ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΙΓΑΙΟΥ
ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ

ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΑΚΩΝ ΚΑΙ ΕΠΙΚΟΙΝΩΝΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ

"ΗΛΕΚΤΡΟΝΙΚΗ ΔΙΑΚΥΒΕΡΝΗΣΗ"

A Survey on Privacy in the Cloud: Threats, Attacks and Countermeasures

Μια επισκόπηση της ιδιωτικότητας στο Νέφος: απειλές, επιθέσεις και αντίμετρα

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

Δρ. Αθανάσιου Γ. Παναγόπουλου (ΑΜ : 3262023015)

Επιβλέπων: Αν. Καθηγητής κ. Κυριάκος Κρητικός

Μέλη εξεταστικής επιτροπής:
Καθηγητής κ. Σπύρος Κοκολάκης
Καθηγήτρια κ. Μαρία Καρύδα

Σάμος, Μάρτιος 2025

Στην αγαπημένη μου σύζυγο Έφη,
που στηρίζει κάθε μου ακαδημαϊκή
προσπάθεια με περηφάνεια και υπομονή

This page was intentionally left as blank

Πρόλογος και ευχαριστίες

Η παρούσα μεταπτυχιακή εργασία στην ασφάλεια και στην ιδιωτικότητα του Υπολογιστικού Νέφους (ΥΝ) (Cloud) καθώς και οι επιθέσεις και οι απειλές, ασχολείται με ένα από τα περισσότερο συζητήσιμα θέματα στον κόσμο της Πληροφορικής και της Ηλεκτρονικής Διακυβέρνησης (Η.Δ.). Το Υπολογιστικό Νέφος (ΥΝ) (Cloud) δεν είναι μια φυσική οντότητα, αλλά αντίθετα είναι ένα τεράστιο δίκτυο απομακρυσμένων διακομιστών σε όλο τον κόσμο που είναι συνδεδεμένοι μεταξύ τους και προορίζονται να λειτουργούν ως ένα ενιαίο οικοσύστημα. Το ΥΝ προσφέρει μεγάλη ευελιξία και αξιοπιστία, αυξημένη απόδοση και αποδοτικότητα καθώς και συμβάλλει στη μείωση του λειτουργικού κόστους των επιχειρήσεων γενικότερα. Βελτιώνει επίσης την καινοτομία, επιτρέποντας στους οργανισμούς να επιτύχουν ταχύτερο χρόνο για την εμπορία και την ενσωμάτωση π.χ. περιπτώσεων χρήσης τεχνητής νοημοσύνης και μηχανικής μάθησης στις στρατηγικές τους.

Στον σύγχρονο και ηλεκτρονικό αυτόν κόσμο δεν θα ήταν εξαίρεση από τις επιθέσεις και τις απειλές και το ΥΝ. Η δε έρευνα για την έγκαιρη αντιμετώπιση τους είναι ερευνητικό σημείο αιχμής. Η Η.Δ. ιδιαίτερα στην Δημόσια Διοίκηση, έχει ως μια εκ των παραμέτρων της όχι μόνο την χρήση του ΥΝ (και της νεφο-υπολογιστικής, cloud computing) αλλά την αντιμετώπιση των σχετικών επιθέσεων. Ιδιαίτερα στις μέρες μας, όπου η χρήση εξαιρετικού όγκου δεδομένων, η ανάγκη υψηλής και κατανεμημένης υπολογιστικής ισχύος και αποθήκευσης δεδομένων αλλά και η τήρηση της ιδιωτικότητας τους, καθιστούν αναγκαία και γεμάτη προκλήσεις μια τέτοια έρευνα. Μια έρευνα, που με την τιμή που μου εδόθη ως μεταπτυχιακός φοιτητής στο συγκεκριμένο ΠΜΣ της ΗΔ του Παν/μιου Αιγαίου, θεώρησα ως σημαντική να παρέχω με εμβρίθεια, όντας και εγώ επαγγελματικό μέλος της δημόσιας κοινότητας όπου έρχομαι αντιμέτωπος πολλές φορές με την ανάγκη προστασίας των δεδομένων στο ΥΝ της ΔΥΠΑ, όπου και εργάζομαι.

Στην ερευνητική αυτή προσπάθεια και συμμετοχή στο ΠΜΣ, έναυσμα ήταν η απότομη τεχνολογική χρήση της Πληροφορικής τα τελευταία χρόνια στην ΔΥΠΑ και στην Ελληνική Δημόσια Διοίκηση, ιδιαίτερα σε εφαρμογές στο ΥΝ, η ερευνητική μου διάθεση για εμβάθυνση σε θέματα ασφάλειας της πληροφορίας στο ΥΝ, καθώς και η φιλοδοξία για την κατάθεση μιας ερευνητικής προσπάθειας, που θα συγκεντρωθεί στα είδη των επιθέσεων και τα αντίμετρα στο ΥΝ, ως αναφορά και επισκόπηση. Σε αυτήν δε την προσπάθεια, ίσταται όπως και έως σήμερα, η οικογένεια μου, σε όλες τις ακαδημαϊκές μου διαδρομές, αλλά και η υπηρεσία που εργάζομαι διότι το άριστο περιβάλλον της είναι το καθημερινό μου κίνητρο για έρευνα και συμβολή στις πληροφοριακές της υποδομές και ανάγκες. Τέλος, θέλω να ευχαριστήσω θερμά τον επιβλέποντα της εργασίας, αν. Καθηγητή του Πανεπιστημίου Αιγαίου κ. Κυριάκο Κρητικό, που μου εμπιστεύτηκε την παρούσα εργασία και με καθοδήγησε άοκνα με την επιστημοσύνη του, στην ολοκλήρωση της εκπόνησης της.

Δρ. Αθανάσιος Παναγόπουλος
Τμήμα Μηχανικών Πληροφοριακών και Επικοινωνιακών Συστημάτων
ΠΜΣ : ΗΛΕΚΤΡΟΝΙΚΗ ΔΙΑΚΥΒΕΡΝΗΣΗ
ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΙΓΑΙΟΥ

This page was intentionally left as blank

Table of contents

List of Figures

List of tables

Abbreviations

API	Application Programming Interface
DoS	Denial of Service
ENISA	European Union Agency for Cybersecurity
IaaS	Infrastructure as a Service
IDS	Intrusion Detection System
IEEE	Institute of Electrical and Electronics Engineers
LLM/SLM	Large Language Model / Small Language Model
MITM	Man-in-the-Middle attack
NIST	National Institute of Standards and Technology
PaaS	Platform as s Service
SaaS	Software as a Service
SOAP	Simple Object Access Protocol
TTP	Trusted Third Party
ΤΠΕ	Τεχνολογίες Πληροφορικής και Επικοινωνιών

Περίληψη

Το Υπολογιστικό Νέφος (Cloud Computing) έχει τη δυνατότητα να εξαλείψει τις απαιτήσεις για τη δημιουργία υπολογιστικής υποδομής υψηλού κόστους για τις λύσεις και τις υπηρεσίες που βασίζονται σε ΤΠΕ υποδομές που χρησιμοποιούν οι διάφοροι κλάδοι των επιχειρήσεων. Υπόσχεται να παρέχει μια ευέλικτη αρχιτεκτονική που εφαρμόζεται από τις διάφορες εφαρμογές πληροφορικής, προσβάσιμη μέσω Διαδικτύου για οποιοδήποτε είδος συσκευών.

Αυτό θα επέτρεπε μια πολλαπλάσια αύξηση της χωρητικότητας ή των δυνατοτήτων του υπάρχοντος και του νέου λογισμικού, μέσω της απομακρυσμένης διάθεσης πληθώρας πόρων που διατίθεται από τον πάροχο της υποδομής του Υπολογιστικού Νέφους (ΥΝ). Σε ένα περιβάλλον υπολογιστικού νέφους, όλα τα δεδομένα βρίσκονται σε ένα σύνολο δικτυωμένων πόρων, όπου η πρόσβαση σε αυτά επιτρέπεται μέσω εικονικών μηχανών ή υπηρεσιών αποθήκευσης ανάλογα με το είδος των δεδομένων και τον τρόπο που έχουν αποθηκευθεί. Αναλόγως, και η επεξεργασία των δεδομένων μπορεί να γίνεται από υπηρεσίες ή λογισμικό νέφους. Δεδομένου ότι αυτά τα κέντρα δεδομένων στα οποία στηρίζεται το ΥΝ για την λειτουργία του μπορεί να βρίσκονται σε οποιαδήποτε γωνιά του κόσμου πέρα από την προσιτότητα και τον έλεγχο των χρηστών, υπάρχουν πολλές προκλήσεις ασφάλειας και απορρήτου που πρέπει να κατανοηθούν και να ληφθούν υπόψη.

Η παρούσα διπλωματική εργασία στοχεύει να αναδείξει, επεξεργαστεί, να παρουσιάσει και να αναλύσει τα πολυάριθμα ανεπίλυτα ζητήματα της ασφάλειας και της ιδιωτικότητας που απειλούν την υιοθέτηση και τη διάχυση του ΥΝ και που επηρεάζουν τα διάφορα ενδιαφερόμενα μέρη που συνδέονται με αυτό και συσχετίζονται με την μεγάλη επιφάνεια επιθέσεων και του τεράστιου όγκου δεδομένων που μεταφέρονται στο ΥΝ. Με την παρούσα εργασία θα προσφερθεί μια ολοκληρωμένη και συνοπτική καταγραφή με εύληπτο τρόπο των υπαρχόντων τρωτοτήτων, απειλών και επιθέσεων έναντι της ιδιωτικότητας του ΥΝ, ενώ ταυτόχρονα θα παρουσιαστούν συγκριτικά και θα αναλυθούν διεξοδικά με βάση συγκεκριμένη κατηγοριοποίηση τα αντίμετρα για τις αντίστοιχες απειλές και επιθέσεις στην ιδιωτικότητα.

Τα πλεονεκτήματα της συνεισφοράς της εργασίας, έγκεινται στην σύγκριση των κατηγοριών αντιμέτρων, βασιζόμενη σε ένα κατάλληλα επιλεγμένο σύνολο από κριτήρια. Αναλόγως της ικανοποίησης τους, θα καταλήξουμε στην σύγκλιση των κατηγοριών αντιμέτρων και σε ποιο προτάσσεται ανάλογα με την στρατηγική προστασίας της υποδομής νέφους και των δεδομένων της εκάστοτε επιχείρησης.

Λέξεις Κλειδιά: *υπολογιστικό νέφος, ιδιωτικότητα, δεδομένα, απειλές, τρωτότητες, επιθέσεις, αντίμετρα, βιβλιογραφική επισκόπηση*

Abstract

Cloud Computing (CC) has the potential to eliminate the requirements for setting up high-cost computing infrastructure for the ICT infrastructure-based solutions and services used by various business sectors. It promises to provide flexible architecture implemented by the various ICT applications, accessible via the Internet for any kind of device.

This would allow for a multi-fold increase in the capacity or capabilities of existing and new software, through the remote provisioning of a wealth of resources made available by the Cloud Computing (CC) infrastructure provider. In a CC environment, all data resides in a set of networked resources, where access to it is allowed through virtual machines or storage services depending on the type of data and how it is stored. Accordingly, the processing of data can be conducted by cloud services or software. Since these data centers on which CC relies for its operation may be in any corner of the world, beyond the reach and control of users, there are many relevant security and privacy challenges that must be understood and considered.

This thesis aims to highlight, elaborate, present and analyze the numerous unresolved issues of security and privacy that threaten the adoption and diffusion of CC and that affect the various stakeholders connected with this and correlate with the large attack surface and the huge amount of data transferred to the CC. The thesis will offer a complete, concise and in an understandable way a full record of the existing vulnerabilities, threats and attacks against the privacy of the CC, while at the same time the countermeasures for the corresponding threats and attacks will be comparatively presented and thoroughly analyzed according to a specific classification.

The advantages of the thesis contribution lie in comparison of the countermeasure categories based on an appropriately selected set of criteria. Depending on the satisfaction of these criteria, we will arrive at the convergence of the categories of countermeasures and which countermeasure must be selected according to the strategy of protecting the cloud infrastructure and the data of each company.

Keywords: *cloud, privacy, data, threats, vulnerabilities, attacks, protection measures, countermeasures, review*

1 Introduction

In this introductory chapter, we supply a general presentation of cloud computing (CC), of the cloud privacy and security issues that exist, as well as the benefits that arise by protecting CC from attacks and why cloud privacy is so important. An analysis of the subject, the main contribution and the structure of the thesis then follows.

1.1 Privacy in the cloud

CC concerns the on-demand availability of computational resources, especially data storage (cloud storage) and computing power, without direct active management by the user. On the other hand, the actual CC offerings are virtual resources (like virtual machines – VMs) that can be managed (e.g., VMs are created, started, stopped, destroyed) by the user in the context of their consumption and usage.

Large clouds often operate in a distributed manner across multiple locations, each of which is typically a data center. CC relies on the sharing of resources to offer cloud services on top of them and typically uses a pay-as-you-go model, which can help in reducing capital and operational expenses, although some operational expenses could be regarded as unexpected by inexperienced customers in the CC usage. This model focuses mainly on certain resources, like VMs, and does also consider other accompanying resources like storage and data.

Today, we are living in the era of big data, with companies generating, collecting, and storing vast amounts of data by the second, ranging from highly confidential business or personal customer data to less sensitive data like behavioral and market analytics. Beyond the growing volumes of data that companies need to be able to access, manage, and analyze, organizations are adopting cloud services to help them achieve more agility and faster times to market, and to support increasingly remote or hybrid workforces.

Cloud privacy is the concept of keeping the clients’ private data in the cloud safe from any kind of unauthorized action that could jeopardise their integrity, confidentiality and availability. Cloud customers entrust their sensitive data to third-party providers, who may not have adequate measures to protect it from unauthorized access, breaches, or leaks. Another challenge of cloud computing is reduced visibility and control. Cloud customers may not have full insight into how their cloud resources are managed, configured, or optimized by their providers. They may also have limited ability to customize or modify their cloud services according to their specific needs or preferences. Further, the location of data might also be uncontrollable, depending on the kind of cloud service exploited. This could lead to the violation of legislation and laws regarding data privacy like GDPR one.

Cloud privacy is strongly related to issues such as privacy attacks. There are many privacy-related vulnerabilities in CC that can be exploited by a wealth of privacy-based attacks, such as SQL injection, Cross Site Scripting (XSS), DoS and DDoS attacks. Although there can be countermeasures to confront such attacks, these countermeasures can vary based on various factors and thus their choice is difficult while it also depends on the customer’s budget, security & privacy requirements and goals as well as its internal security expertise (especially in terms of their installation, configuration and proper use). Cloud computing is related to characteristics, deployment, service models, and attacks, such as Google Hacking, and Forced Hacking, among others, which influences also cloud security.

Setting cloud security and privacy as a top priority is crucial for safeguarding a company and preserving clientele. Customers can build a safe cloud environment that supports the expansion of their business by being proactive and implementing measures like encrypting data, monitoring settings, putting in place robust access restrictions, and routinely educating their staff.

Obviously, a reduced level of security in the cloud opens opportunities for third parties (e.g., provider’s collaborators) and attackers in not only reading but also modifying or deleting cloud data. Thus, the privacy of data can be violated by multiple different threat actors.

Additionally, *cloud data security* is essentially the practice of protecting data and other digital information assets from security threats, human error, and insider threats. It leverages technology, policies, and processes to keep your data confidential and still accessible to those who need it in cloud-based environments.

We can refer now to some of the benefits of cloud data security, as follows:

Greater visibility. Sophisticated cloud data security measures allow us to maintain visibility into the inner workings of the cloud, namely what are the data assets and where they live, who is using the customer’s cloud services, and the kind of data they are accessing.

Easy backups and recovery. Cloud data security can offer several solutions and features to help automate and standardize backups, freeing teams from conducting and monitoring manual backups plus troubleshooting problems. Cloud-based disaster recovery also enables restoring and recovering data and applications in minutes.

Cloud data compliance. Robust cloud data security programs are designed to meet compliance obligations, including knowing where data is stored, who can access it, how it is processed, and how it is protected. Cloud data loss prevention (DLP) can help customers easily discover, classify, and de-identify sensitive data to reduce the risk of violations.

Data encryption. Organizations need to be able to protect sensitive data whenever and wherever it goes. Cloud service providers help them to tackle secure cloud data transfer, storage, and sharing by implementing several layers of advanced encryption for securing cloud data, both in transit and at rest.

Lower costs. Cloud data security reduces total cost of ownership (TCO) and the administrative and management burden of cloud data security. In addition, cloud providers offer the latest security features and tools, making it easier for security professionals to perform their jobs with automation, streamlined integration, and continuous alerting.

Advanced incident detection and response. An advantage of cloud data security is that providers invest in cutting-edge AI technologies and built-in security analytics that help customers to automatically scan for suspicious activities so as to quickly identify and respond to security incidents.

There is an increased attack surface of the cloud, various vulnerabilities that might exist and numerous attacks that can be conducted. Moreover, there are issues, because: (i) security and privacy is a shared responsibility between CSP and customer and (ii) while countermeasures exist, it is difficult to select and combine them for various reasons.

Cloud Service Providers (CSPs) are responsible for selected parts of cloud security and not for its entirety. By following the shared responsibility model, cloud security is a joint effort by both CSPs and customers as they handle different but relevant pieces of cloud security.

For instance, in the IaaS model, the CSP is responsible for the security of the physical resources, but the customer is responsible for the security of the VMs on top of these resources, especially as he/she installs software and moves data in them so as to be able to execute his/her applications. This means that the client’s security tools, measures and architecture should be compatible with those of the providers.

Today, it is "the cloud" that relieves businesses from all these expenses and efforts in technology investment. Cloud service providers offer various types of software, platforms, infrastructure, and storage at very affordable and flexible prices in each field. Data storage outsourcing covers all hardware and personnel costs, inclusive of comprehensive problem-solving and automatic scaling of resources. Partnerships with CSPs allow firms to get back precious time, energy, and capital.

As the organization moves to cloud, reassessment of various strategies and tools becomes necessary to ensure that these meet demands for protection in a cloud environment. Moreover, cloud privacy policies and requirements are the number one targets to satisfy, and this requires assessing the ability of a cloud provider and his/her environment to satisfy them. Also, certifications could play a significant role as well as any kind of documentation from the cloud provider concerning best security practices, measures and architecture that he/she applies to his/her environment.

IT executives were initially resistant to partnering with CSPs, as they were unwilling to entrust the responsibility of controlling their data to third parties. Unbeknownst to many, over the past few years, such sentiment has changed. Today, cloud services represent the fastest-growing segment of IT spending, for which the industry was valued at $312B back in 2019, as it is showing over 15 percent y-o-y growth. This can be attributed to the following fact. CSPs are fantastic safety partners because they have the money to hire professional technology talent and house optimum security technologies for their networks. Most often, they offer a full suite of security tools. These include the latest anti-malware software, intrusion protection systems, application firewalls, network monitoring, and event analysis solutions. These tools run on machine learning as well as artificial intelligence. CSPs are also better equipped to counter distributed denial-of-service attacks than most companies.

However, CSP's best-practice security enclave comes at a time when DDoS attacks rose by 242% in the first quarter of 2020, and 542%[1] over the preceding quarter. Infrastructure security is a concern in cloud computing, but cloud privacy is the major concern. Especially, as new regulations regarding privacy, evolving cybersecurity threats, and recent data breaches leading to some serious privacy class action lawsuits have only increased this worry. One of the major areas of concern as far as security is concerned involves compliance with international data laws. These are mostly concerned with the aspect of privacy and any other personal information of citizens.

Many organizations remain hesitant to migrate sensitive data to the cloud as they struggle to understand the cloud providers, their security options and meet regulatory demands.

The traditional network perimeter is fast disappearing, and security teams are realizing that they need to rethink current and past approaches when it comes to securing cloud data. With data and applications no longer living inside a company's data center and more people than ever working outside a physical office, companies must solve how to protect data and manage access to that data as it moves across and through multiple environments, including CC ones.

1.2 Thesis subject

The selection of the right combination of complementary countermeasures is a difficult task to perform which depends on too many factors, including the privacy and security requirements of a business. Even the installation, configuration and execution of such countermeasures can be a rather complex task that requires deep security expertise, which might not be available in a company/organization.

The final goal of the thesis and its main contribution is to provide a useful research approach in confronting cloud threats and attacks and to compare proposed countermeasures to finally select the most proper for the organization at hand, under specific evaluation criteria and based on an organization's strategy, which, in turn, is the main advantage for the benefit for the reader (as a potential organization representative).

More specific, the thesis contributes to cloud privacy by analyzing the relevant attacks, threats and vulnerabilities, presenting a taxonomy of suitable countermeasures, developing a set of criteria to classify the countermeasures categories, study the convergence and the best category depending on the criteria importance. A rigorous analysis will be conducted for the root causes and potential impacts of these attacks on sensitive data stored and processed in cloud environments.

The thesis also investigates innovative privacy-enhancing technologies, such as advanced encryption schemes, homomorphic encryption or Privacy-Enhancing Technologies to mitigate the identified threats. Rigorously evaluating performance, security, and privacy guarantees these techniques in realistic cloud settings.

[1] Ellen, Daniel, "*DDoS attacks rose by 542% in Q1 of 2020*", Verdict report, 21 August 2020, https://www.verdict.co.uk/ddos-attacks/?cf-view

Then, the countermeasures will be classified into various categories (DLPS, Data storage and Privacy as a Service, MiTM, ISD and Data privacy protection mechanisms), evaluated by criteria and tagged accordingly (with evaluation values as Low, Moderate, High) while we conclude with the countermeasures category(ies) which satisfy(ies) most of the criteria.

Additionally, in the thesis we will analyze every such countermeasure category for each one of the selected criteria depending on the cloud environment of the organization and we will present a grip, innovative and structured methodology and mainly an *algorithm* to conduct results on the best countermeasures as a guidance for an organization to benefit by these and built its IT strategy against cloud attacks protecting cloud privacy. This will supply proper guidance for organizations so as to benefit by building their IT strategy for enhancing cloud privacy and better protecting their data and workloads.

Finally, the thesis will also contribute to the research by presenting the future and emerging attacks, threats and vulnerabilities by paying attention to emerging threats using a user-centric approach, prioritizing user privacy and control in cloud that empower users to manage their data and privacy settings effectively.

Academics, researchers and security professionals will benefit by the thesis, as they will find a rigorous and grip presentation of the attacks affecting cloud data privacy, a presentation and classification of the attacks under categories and a professional guide with innovative privacy-enhancing technologies, with the classification of the most known countermeasures of every category of attack. Readers will find a set of criteria for using the best-fit countermeasures, based on a structured methodology and an innovative algorithm. Under this algorithm, they can realize that using the converged criteria, what are the proposed countermeasures which could be used to protect organization’s cloud data privacy.

Furthermore, in this thesis, we are essentially going to discuss the issue of cloud privacy under the specific reasoning of protecting data and supplying a well-described taxonomy of vulnerabilities and attacks regarding cloud privacy. References to real and important cloud attack cases in the past and their consequences will also be an important part of the thesis. We will also contribute with a well-presented taxonomy, analysis, and comparison of the approaches for countermeasures against the attacks. The current challenges and promising research directions to confront them in terms of cloud privacy further enrich our thesis contribution.

1.3 Thesis structure

This thesis is structured as a research paper, as follows:

Chapter 2 is related to a general presentation of basic terms and topics regarding security and privacy, such as vulnerability, attacks, and threats. The purpose is to make the reader of this thesis report more familiar with the topics and terminology engaged.

In Chapter 3 we explain the literature review methodology we have followed, covering (a) the relevant research questions to answer and where they are answered in the remaining chapters of this report and (b) the way the available literature was searched and filtered. In addition, we present some statistics and quantitative data of the findings produced.

In Chapter 4, we present taxonomies of vulnerabilities, threats and attacks in cloud security and privacy as well as, and analyse them. Further, we present real cases of cloud attacks in the past and their consequences.

In Chapter 5, a taxonomy of privacy-based countermeasures is supplied followed by an analysis and comparison of these countermeasures based on specific criteria. Further, deep insights from the comparison are supplied, covering, for instance, the best countermeasure per each category. Finally, an algorithm is suggested for the selection of the right countermeasure category based on the target organization’s security requirements and preferences and the type of cloud it intends to adopt.

In Chapter 6, we shortly analyse the remaining challenges in the field of cloud privacy and detail some promising research directions to confront them.

Finally, in Chapter 7 we conclude the thesis by outlining the thesis’ main scientific contribution as well as its main findings. In addition, the most important and pressing challenges are marked down along with the most interesting future research directions to follow to confront them.

2 Background on cloud computing, security & privacy

This chapter aims to present the main concepts and terms related to the cloud and the privacy of data stored in it. Its main goal is to equip the prospective reader with the right knowledge so as to increase his/her level of understanding in terms of the remaining content of this thesis report. To this end, this chapter is organized as follows: the first section focuses on clarifying the main cloud concepts and types of services offered while the second section dives into the main concepts and terms related to cloud security and privacy.

2.1 Cloud computing concepts and services

2.1.1 Cloud service and deployment models

A lot of definitions have been supplied that explain cloud computing (CC). CC has been defined as the new state-of-the-art technology that is able to provide a flexible and shareable third-party IT infrastructure that the businesses can exploit in order to dynamically provision their services and business processes in an elastic manner. This integrates features supporting high scalability and multi-tenancy (B. R. Kandukuri, et al., 2009).

NIST defines “Cloud computing as a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model is composed of five essential characteristics, three service models, and four deployment models” (see Fig. 1).

CC has gained wide acceptance for organizations as well as individuals by proposing computation, storage and software-based services. It is used to address the resource scarcity issues of its clients by providing them with on-demand pay-per-use services (Buyya et al., 2011). A *cloud* is a centralized collection of resources connected through a high-speed network. Its global availability of high-performance resources, support for many services, and ability to store a large amount of data have made it ubiquitous. Even with modern smartphones, cloud computing is able to serve multiple purposes ranging from the backup of contacts to the execution of complex applications through computation offloading (Sanaeiet al. 2014, Kumaret al. 2013). Moreover, services at relatively low cost and with an assurance about the quality of services make it an

attractive solution for mitigating the resource constraint problem[2]. Since a cloud computing platform provides services through the sharing of valuable resources, adequate usage of these resources may be achieved by ensuring that the platform is able to counter security threats which otherwise may degrade its performance and reliability.

Furthermore, CC reduces capital expenditure as well as operational costs. Generally, the payment is based on a pay-per-use and subscription model. Moreover, CC eliminates the need to install and run a cloud application offered on the customer's local computer, thus alleviating the customer's burden on software installation and maintenance. In fact, all operational aspects of applications are taken care of by the cloud (application) provider.

Based on the various types of services provided, CC can be said to have three layers or delivery models (see Fig. 1):

1) IaaS - Infrastructure as a Service (IaaS). The capability offered to the consumer relates to the provisioning and aggregating processing, storage and other fundamental computing resources in the form of virtual machines in which the consumer is able to deploy and run arbitrary software that can include operating systems (OSs) and applications. The consumer does not manage or control the underlying cloud infrastructure but has control over all the elements that constitute the content of a VM, such as OSs, data, and deployed applications, as well as over the VM itself. Further, the consumer has possibly limited control of select networking components (e.g., firewalls, dynamic IPs, Virtual Private Networks - VPNs, etc.). Finally, he/she can control the virtual storage attached to the provisioned VMs. The main advantage of IaaS is flexibility in terms of both the VM capabilities as well as the content to be deployed on them plus the convenient pay-per-use charge model. Well-known IaaS services include AWS EC2, Google Cloud Platform (GCP) Compute Engine and Microsoft Azure VMs.

2) PaaS – Platform as a Service (PaaS). The capability offered to the consumer is to deploy a dedicated platform onto the cloud infrastructure for developing and executing consumer applications. For that reason, the platform encompasses relevant development and runtime features, such as programming language runtimes, libraries, services, frameworks, middleware and tools. Many PaaS services also offer the capability to scale the offered platforms on demand depending on customer-supplied conditions, supplying also a load-balancing service on top of the platform instances provisioned. The consumer does not manage or control the underlying cloud infrastructure, including network, servers, operating systems, and storage, and the provisioned VM on which the dedicated platform operates. However, the consumer has control over the deployed applications and data as well as possibly configuration settings for the platform’s environment. Generally, PaaS works well for small businesses and startup companies for two very basic reasons. First, it is cost effective, allowing smaller organisations access to state-of-the-art resources without a big price tag. Most small firms have never been able to build robust development environments on premise, so PaaS provides a path for accelerating software development. Second, it allows companies to focus on what they specialize in (i.e., application development) without worrying

[2] A resource constraint problem deals with scenarios where personnel or workforce employed to perform the tasks are limited and each job has an arrival time, a due date, and a penalty associated to delays.

about maintaining basic infrastructure and development platform. Some well-known PaaS services include the Google App Engine, Microsoft Azure App Service, AWS Beanstalk, Red Hat OpenShift and IBM Cloud Foundry.

3) SaaS – Software as a Service (SaaS). The capability offered to the consumer is actually a whole application running on a cloud infrastructure. The applications are accessible from various client devices through either a thin client interface, such as a web browser (e.g., web-based email client application), or a program interface (e.g., CLI, RESTful API, etc.). The consumer does not manage or control the underlying cloud infrastructure and platform or even individual application capabilities, except for limited user-specific application configuration settings. In the previous delivery models, the customer is responsible for his/her data and has full control over them. However, in the SaaS model, the control over customer is much limited.

In this model, the SaaS provider builds the core application structure, ensuring a strong foundation (physical, application, and network security). In SaaS, two similar cohesion mechanisms can be applied. In the first model, two types of servers are used: the Main Consistence Server (MCS) and Domain Consistence Server (DCS). Cache coherence in SaaS is achieved by the cooperation between MCS and DCS. However, if the MCS is damaged or compromised, control over the cloud environment will be lost. Therefore, it is of great concern to fully secure the MCS, so as to also secure control over the cloud environment from the side of the application provider. SaaS providers do employ a cohesion mechanism which has some similarities with the MCS/DCS one. As an example, Amazon RDS (Relational Database Service) can be configured in a multi-region setup with a primary region (MCS-like) ensuring strong consistency, while read replicas (DCS-like) in other regions provide localized performance and eventually sync with the primary region. The second cohesion mechanism is called Divided Cloud and Convergence which shares similar scope, goals and principles with the MCS-DCS model.

The main advantages of the SaaS delivery model are flexible charge models, no application installation and maintenance cost, ubiquitous access and high application availability. Some well-known SaaS services include HubSpot, Dropbox, Zoom, Slack and Shopify.

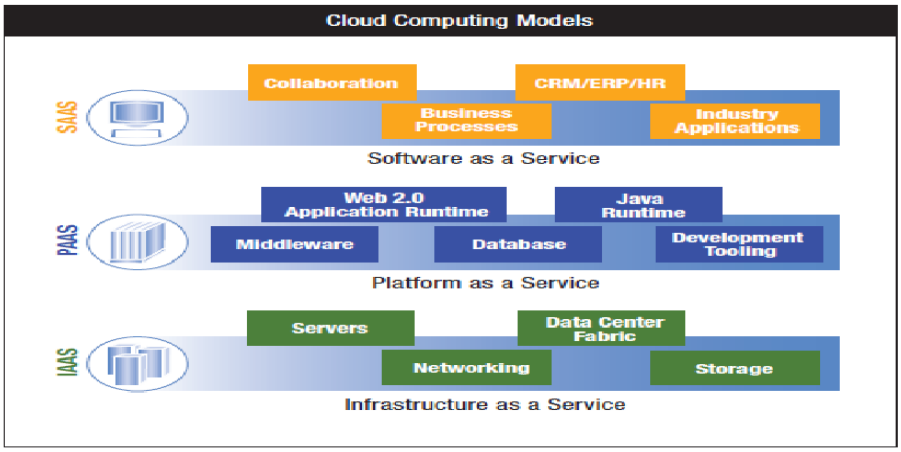

Fig. 1: The cloud delivery models

The five essential cloud characteristics according to the NIST definition are as follows:

On-demand self-service. A consumer can unilaterally request and automatically achieve the provisioning of computing capabilities, such as server time and network storage, on-demand without requiring human interaction with each service provider.

Ubiquitous access. The offered cloud services are available over the network/Internet and accessed through standard mechanisms that promote the use by heterogeneous thin or thick client platforms (e.g., mobile phones, tablets, laptops, and workstations). These services are continuously supplied and can be consumed in any place (i.e., from all over the world).

Multi-tenancy. The provider’s computing resources are pooled to serve multiple consumers using a multi-tenant model. In such a model, it is possible that physical resources are shared between clients. This can occur when on top of the same physical resource different virtual resources are assigned to different customers. In this sense, with CC environments different physical and virtual resources might be dynamically assigned and reassigned according to consumer demand.

Multi-tenancy originally mapped to an architecture in which a single instance of a software application serves multiple customers. So, through CC, this term has been extended to cover the sharing of any kind of computational resource.

In CC there is a sense of location independence in that the customer generally has no control or knowledge over the exact location of the resources provided but may be able to specify the desired location at a higher level of abstraction (e.g., country, state, or datacenter). Examples of (multi-tenant cloud) resources include storage, processing, memory, and network bandwidth.

Rapid elasticity. Offered capabilities and services can be elastically provisioned and released, in some cases automatically, to scale rapidly outward and inward commensurate with demand. To the consumer, the capabilities available for provisioning often appear to be unlimited, especially in the context of cloud providers, and can be appropriate in any quantity at any time.

Measured service. Cloud systems automatically control and optimize resource use by leveraging a metering capability at some level of abstraction appropriate to the type of service (e.g., storage, processing, bandwidth, and active user accounts). Resource usage can be monitored, controlled, and reported, providing transparency for both the provider and consumer of the utilized service[3].

Irrespective of the above-mentioned service models, clouds (services) can be deployed in four different ways depending upon who owns and who consumes them (Deployment models’ above) (Khan, 2016). These deployment models are shortly analysed as follows:

a. *Private cloud.* The cloud infrastructure is provisioned for exclusive use by a single organization comprising multiple (internal) consumers (e.g., business units or departments). It may be owned, managed, and operated by the organization itself, a third party, or some combination of them. It may also exist on (e.g., in own data centre) or off (e.g., in the third party’s data centre) premises.

[3] Cloud services generally charge users per hour of resource usage or based on the number of certain kinds of transactions that have occurred, the amount of storage in use, and the amount of data transferred over a network. All usage is measured.

As private clouds are usually located in privately hosted environments, users are afforded more control over their data, applications and overall infrastructure, and benefit from on-demand resources. This gives businesses the ability to handle small and medium-sized workloads and flexibly add more computer power when needed. In this context, all offered resources (e.g., VMs) become available over a proprietary (organizational/enterprise) network. In this way, a private cloud can make it easier for an organization to customize its resources to meet specific IT requirements.

Private clouds are often used by government agencies, financial institutions, and other mid- to large-size organizations with business-critical operations seeking enhanced control over their environment and a high-level security for their data and digital assets in general.

Advantages of a private cloud include:

- **More flexibility** – an organization can customize its cloud environment to meet specific business needs (e.g., security ones).
- **More control** – resources are not shared with others, so higher levels of control, security and privacy are possible.
- **More scalability** – private clouds often offer more scalability compared to on-premises (non-cloud) infrastructure. However, this level of scalability is smaller with respect to other cloud deployment models / types.

As disadvantages of the Private Cloud, we can indicate that a private cloud can be difficult for organizations to maintain on their own due to tasks like performance monitoring, software patching and planning updates. Moreover, private cloud deployments require high levels of technical expertise.

b. *Community cloud.* The cloud infrastructure is provisioned for exclusive use by a specific community of consumers from organizations that have shared concerns (e.g., mission, security requirements, policy, and compliance considerations). It may be owned, managed, and operated by one or more of the organizations in the community, a third party, or some combination of them. As in the case of a private cloud, a community cloud may exist on or off premises (so either in the premise of the community data centres or at the data centre(s) of a third-party).

The benefit of the community cloud model is mainly cost efficiency. Sharing infrastructure reduces individual capital and operational expenditure. In particular, organisations benefit from economies of scale in purchasing and maintaining hardware and software. Further, the management and maintenance effort and cost can be shared between the community members/organisations. Science clouds are an interesting example of community clouds. In this case, the common interest driving different scientific or research organizations sharing a large, distributed infrastructure is scientific computing. Finally, a community cloud is more secure than a public cloud (as external access is forbidden or quite limited) but potentially less secure than a private cloud (as internal threats are more probable).

As disadvantages of the Community Cloud, we can indicate that if many users search for something in your system at the same time, it can lead to longer loading times, in comparison to private cloud (because usually a community cloud is larger than a private cloud[4]). You need to

[4] In the context of search, data can be replicated and search could be load-balanced. In this context, search performance might be better than in the private cloud.

purchase a license for software like OS, for each user. This can become expensive with many users. However, the prices per user usually decrease significantly (although is sometimes shared between the community members). Additionally, security is less than in a private cloud. Scalability is also smaller than a public cloud. We also have more complexity, especially in maintenance with respect to a private cloud.

c. *Public cloud.* The cloud infrastructure is provisioned for open use by the general public. It may be owned, managed, and operated by a business, academic, or government organization, or some combination of them. It (usually) exists on the premises of the cloud provider. The services offered are available for consumption by the general public usually through the Internet based on the aforementioned pay-per-use model through a multi-tenant manner. Usually, public cloud infrastructures are large and able to accommodate workloads of any kind and size. Public cloud is the most common cloud deployment model. Microsoft Azure and Amazon AWS are examples of well-known public clouds.

Advantages of public clouds:

- **Lower costs –** the customers do not need to purchase hardware or software while they pay only for the services that they use.
- **No maintenance –** the cloud provider is responsible for the operation and maintenance of the infrastructure as well as of the services and applications that he/she is offering on top of it.
- **Near-unlimited scalability –** on-demand resources are available at any time to meet your business needs. The resources are numerous, giving the illusion of infinite scalability, especially in the context of big cloud providers.
- **High reliability –** a vast network of servers within the cloud infrastructure with appropriate fail-over mechanisms ensures against failure.

d. *Hybrid cloud*. The cloud infrastructure is 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 (e.g., cloud bursting for load balancing between clouds). A typical form of a hybrid cloud involves the combination of on-premises infrastructure —or a private cloud — with a public cloud. In this way, hybrid clouds allow data and apps to move between the two environments, enabling the realization of cloud bursting scenarios, where data and computation move or is replicated to the public cloud when the capability of the private cloud becomes fully occupied/reserved.

Many organizations choose a hybrid cloud approach due to business imperatives, such as meeting regulatory and data sovereignty requirements while still maintaining satisfactory levels of performance and availability. For instance, a hybrid cloud enables us to maintain private/sensitive data in the private cloud part (which is more secure) while other non-sensitive data plus computation can be moved, processed and stored in the public cloud part (which is more scalable and performant). In case private data needs to be processed, then either the computation is immediately done at the private cloud part, or the data is moved in a secure manner to the public cloud part (e.g., through encryption) so as to be processed there. In the latter case, once computation is over, the data is erased to reduce the risk of being compromised due to any kind of security issue.

The hybrid cloud is evolving to include edge workloads as well. Edge computing brings the computing power of the cloud to IoT devices, i.e., closer to where the data resides. By moving workloads to the edge, devices spend less time communicating with the cloud, reducing latency and energy consumption.

A hybrid cloud platform gives organizations many advantages—such as greater flexibility, more deployment options, security, compliance, and getting more value from their existing infrastructure. When computing and processing demand fluctuates, hybrid cloud computing gives businesses the ability to seamlessly scale up their on-premises infrastructure to the public cloud to handle any overflow—without giving third-party datacenters access to the entirety of their data. Organizations gain the flexibility and innovation the public cloud provides by running certain workloads in the cloud while keeping highly sensitive data in their own datacenter to meet client needs or regulatory requirements.

This not only allows companies to scale computing resources— it also eliminates the need to make massive capital expenditures to handle short-term spikes in demand, as well as when the business needs to free up local resources for more sensitive data or applications. Companies will pay only for resources they temporarily use instead of having to purchase, program, and maintain additional resources and equipment that could remain idle over long periods of time.

e. Multi-Cloud.

A multi-cloud environment is an IT strategy (not actually a cloud development model) where organizations use multiple cloud computing services from different providers. This approach helps mitigate risks, enhance performance, and optimize costs by distributing workloads across various cloud platforms, such as AWS, Microsoft Azure, Google Cloud Platform (GCP), and others.

Benefits of Multi-Cloud Environments, are consider the following:

- 1. Avoiding Vendor Lock-in. Organizations gain flexibility by not being tied to a single provider. Also, it enables seamless migration of workloads between cloud platforms.
- 2. Improved Performance and Redundancy. Multi-cloud architecture ensures higher availability and disaster recovery and distributes workloads across clouds to optimize performance and reduce latency.
- 3. Cost Optimization. Enterprises can compare pricing models and choose the most cost-effective options, and it reduces reliance on a single cloud provider’s pricing structure.
- 4. Enhanced Security and Compliance. Meets regulatory compliance by storing sensitive data in specific regions and it reduces risk exposure by segmenting workloads across different clouds.
- 5. Best-in-Class Services. Organizations can leverage specialized services from different providers (e.g., AI/ML from Google Cloud, enterprise applications from Microsoft Azure, and compute services from AWS).

As challenges of Multi-Cloud Environments, we can refer:

- 1. Complexity in Management. Managing multiple cloud platforms requires specialized tools and expertise and while ensuring seamless integration between different cloud providers can be challenging.

- 2. Security and Compliance Risks. Security policies must be standardized across providers to prevent misconfigurations. Data sovereignty and compliance regulations vary across cloud providers.
- 3. Data Transfer Costs. Moving data between different clouds may incur high costs and organizations need to optimize data transfer strategies to reduce expenses.
- 4. Performance Monitoring and Optimization. Ensuring consistent performance across multiple cloud providers requires advanced monitoring tools and network latency between cloud providers can affect performance.

For the best practices for Multi-Cloud Deployment, we note:

- 1. Implement Cloud-Native Security Solutions. Use Cloud Access Security Brokers (CASBs) for visibility and control and employ Zero Trust security models to enhance identity and access management.
- 2. Standardize Policies and Compliance Controls. Utilize Cloud Security Posture Management (CSPM) tools to ensure compliance and automate security audits and monitoring.
- 3. Optimize Workload Distribution. Use load balancing to distribute workloads effectively across cloud providers and deploy Kubernetes or container orchestration tools for workload portability.
- 4. Leverage Multi-Cloud Management Tools. Use platforms like Google Anthos, AWS Outposts, and Azure Arc for centralized management and implement Infrastructure-as-Code (IaC) tools, such as Terraform, for automation.
- 5. Monitor and Optimize Costs. Utilize cloud cost management tools to analyze spending patterns and implement auto-scaling and reserved instances to optimize expenses.

A multi-cloud environment provides flexibility, resilience, and cost savings but comes with management complexities. Organizations must adopt best practices, leverage automation, and implement strong security controls to fully benefit from multi-cloud strategies. By doing so, they can achieve greater business agility, improved security posture, and optimal cloud resource utilization.

2.1.2 Cloud configuration

Understanding cloud security issues requires, initially, to realize the cloud configuration that we will discuss in the following paragraphs. A cloud company is made up of resources dedicated to satisfying respective demands. According to NIST, the cloud computing configuration (as a high-level architecture of the whole cloud ecosystem) has five major actors enlisted in the following figure (Fig. 2). We can also refer to the CC configuration as a *cloud ecosystem architecture which is defined as a complex system of cloud services, platforms, and infrastructure used for the storage, processing, and distribution of data and applications through the Internet.* It consists of multiple parts: cloud providers, software developers, users, and other services, which are integrated into a prolific and adaptable architecture for computing assets. This ecosystem enhances the ability of businesses and individuals to lease computational solutions at will, in line with flexibility,

innovation and cost sensitivity in the digital frontier. The obligations or duties of each role from these five are analyzed as follows:

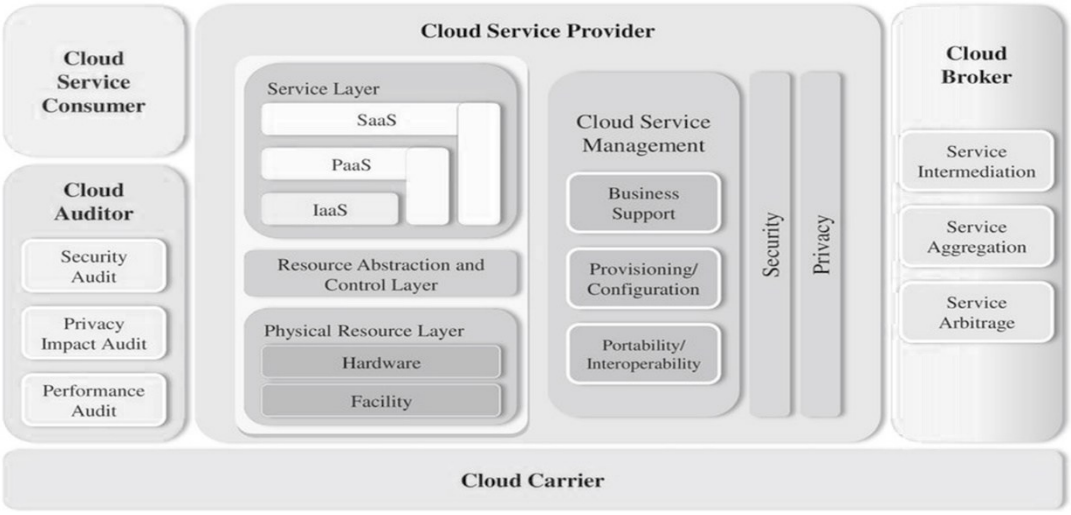

Fig. 2 NIST cloud computing configuration

• *Cloud consumer.* A cloud consumer is a person or organization who receives one or more services offered by a cloud provider. A cloud consumer usually may first select the most appropriate service by investigating the services offered by multiple cloud providers and then conclude the respective contract with the provider of the selected service.

• *Cloud service provider*. A cloud provider is a person or organization who makes a cloud service available to a cloud consumer. The cloud service might operate under resources which are either owned by the CSP (so CSP is also a cloud provider) or are offered by a third-party (i.e., a cloud provider). A CSP and a cloud provider are thus two different roles although they can be played by the same (legal) entity.

• *Cloud auditor.* The role of a cloud auditor is to independently audit cloud services. An auditor would review the conformance to standards based on several objective evidence reviews. Cloud auditors may review services provided by the cloud providers in terms of privacy impacts, security controls, performance, etc.

• *Cloud broker*. Because the consumption and integration of cloud services might be too complicated for cloud consumers, they can obtain their cloud services through a cloud broker instead of contacting a cloud provider directly. In fact, the work of a cloud broker is to manage the usage, performance, and delivery of cloud services as well as the relationships among cloud consumers and cloud providers. Cloud broker services can be classified into three categories:

(i) *Service intermediation.* A cloud broker primarily has the capability of enhancing some particular capabilities of certain cloud services and add value to them for the sake of cloud consumers. Examples of such enhancements include access management of cloud services, performance reporting, identity management, enhanced security, etc.

(ii) *Service aggregation:* This cloud broker capability concerns the aggregation or combination of multiple individual cloud services into a one new composite cloud service. As a matter of fact, the cloud broker might be also capable of integrating data (by, e.g., using different data services) and thus play the role of a (data) bridge between cloud consumers and several cloud

providers. Thus, not only software services and applications can be combined but also the data managed or processed by them.

(iii) *Service arbitrage* acts like service aggregation but in a more dynamic manner as the composite services are produced dynamically from the individual cloud services (in contrast to service aggregations which constitute fix service compositions). As a matter of fact, service arbitrage gives flexibility to the broker in selecting services from different cloud providers so as to satisfy the requirements from specific classes of consumers.

• *Cloud carrier*. A cloud carrier acts as an intermediary between the cloud consumer and the cloud provider to guarantee cloud service delivery through network-related capabilities/services. By establishing SLAs with a cloud carrier, the cloud provider can bring services consistent with the offered SLAs to cloud consumers (as network latency is guaranteed by the cloud carrier such that service performance is also reassured).

Additionally, *virtualization* is the foundation of cloud computing, enabling multiple virtual environments to run on a single physical machine. While it provides flexibility, scalability, and cost savings, it also introduces potential attack vectors unique to virtualized environments. (Scarfone, 2011[5]). Virtualization refers to creating virtual instances of computing resources (e.g., servers, storage, networks) using a hypervisor. It allows multiple virtual machines (VMs) or containers to run on shared physical hardware, abstracting the underlying resources from users. In some kinds of virtualizations, it is very easy to share data between the systems. Sometimes this feature can turn into an attack vector if it is not well controlled.

2.2 Cloud security and privacy terminology

With the development of CC and other ICT technologies as well as the increasing accessibility to the Internet, organizations become vulnerable to various types of threats. In fact, their data becomes exposed to cyber-attacks and their resulting damage. *Threats* come from different sources, like employees' activities or hacker attacks. *Vulnerabilities* consist of weaknesses in a system (like a cloud environment) which can be exploited by the attackers that may lead to dangerous impact when their attacks succeed. When vulnerabilities exist in a system, a cloud threat may be manifested via a threat agent using a particular penetration technique to cause undesired effects (Jouini et. al, 2014[6]).

2.2.1 Cloud security

Cloud security refers to the cybersecurity policies, best practices, controls, and technologies used to secure applications, data, and infrastructure in cloud environments. In particular, cloud security strives to provide storage and network protection against internal and external threats, access management, data governance and compliance, and disaster recovery. Security services

[5] Scarfone, K., 2011. "*Guide to Security For Full Virtualization Technologies*",. DIANE Publishing.

[6] Mouna Jouini, Latifa Ben Arfa Rabai, Anis Ben Aissa, "*Classification of security threats in information systems*", 5th International Conference on Ambient Systems, Networks and Technologies (ANT-2014), Procedia Computer Science 32 (2014) 489 – 496, ScienceDirect, Elsevier

covering integrity, confidentiality, authentication, non-repudiation, and availability secure both computer networks and information systems (Tabrizchi and Kuchaki Rafsanjani, 2020[7]). These latter security properties are formally explained as follows:

• *Confidentiality* ensures and means that information should not be disclosed and become available to unauthorized individuals, entities, or processes. As a matter of fact, data should be sent and received without being accessed by unauthorized entities (e.g., intermediaries) during transmission. Further, data should not be accessed by unauthorized entities while at rest. One good way to realize confidentiality is data encryption. Encryption can be performed through the symmetric or asymmetric key paradigm. Another good way, especially in the context of data at rest, are access control policies and respective mechanisms that restrain access to the data based on the requirements of its owners.

• *Integrity* ensures that the information transmitted from its source is received unmodified. It also ensures the validity of data when they are modified while at rest. In other words, it ensures that the information has not been altered by any unauthorised party, either intentionally or unintentionally.

• *Availability* ensures that data is accessible and usable once in demand by an authorized party. For example, when a DDoS attack occurs in a system, the system will lose its ability to avail data over authorized requesters.

• *Authentication* must confirm the identity of the information transmitter and recipient. In fact, the integrity and confidentiality of information are meaningful just when the identities of senders and receivers are properly verified. Traditionally message/data authentication meant confirming the identity of the transmitter. However, in the context of data privacy, this definition needs to be broadened so as to reassure that data ends up to an authorized party/receiver.

• *Non-repudiation:* This ensures that actions taken cannot be denied by senders or receivers. There are basically two non-repudiation types: source repudiation and destination repudiation. In the former, the sender cannot deny the transmission of a message, while in the latter, the recipient cannot deny the delivery of a message.

The cloud platform is usually equipped with high performance server machines, high speed storage devices and an efficient network. Cloud users who have mobile phones, laptops or modern desktops connect to the cloud platform through the Internet. Since the server machines are inter-connected through an internal network, an attack on the network can cause a disastrous effect like communication delays or even inaccessibility of the network (McMillan, 2009, InfoSecurity, 2009). Conclusively, *a cloud attack is a malicious activity that targets systems, services, infrastructure, or data hosted in a cloud*. These attacks aim to compromise cloud resources' integrity, availability, or confidentiality. Cloud attacks include data breaches, data loss, account or service traffic hijacking, insecure interfaces, denial of service (DoS), and more. However, even sniffing can cause private data disclosure when data is transmitted in an unencrypted form. The attacks on virtual machines and hypervisors being used for running the virtual machines have also been proven to cause extensive breaches in security as well as privacy for malicious purposes. Similarly, cloud services

[7] Tabrizch, H., Kuchaki-Rafsanjani, M., "*A survey on security challenges in cloud computing: issues, threats, and solutions*", The Journal of Supercomputing (2020) 76:9493–9532, https://doi.org/10.1007/s11227-020-03213-1, Springer

can also be prone to security threats as this layer contains software that has always been vulnerable to security attacks.

Cloud attacks are cybercriminal actions that often exploit vulnerabilities and weaknesses in cloud security to gain access to valuable data and assets. Once attackers access cloud account credentials, they impersonate legitimate users. This may lead to some types of attacks that could cause a violation of data protection, which can also affect making services unavailable to all clients. Therefore, *cloud vulnerabilities are weaknesses, oversights, or gaps in cloud infrastructure that attackers or unauthorized users can exploit to gain access into an organization's environment and potentially cause harm.*

2.2.2 Cloud security policies and standards

Security policies (Ahmed and Litchfield, 2018[8]) can be defined as *high-level documents that specify an organization's goals, rules, and guidelines for managing and protecting its information and systems*. They can also include reactive measures. In general, they span various areas, such as governance and compliance, risk management, access control, awareness and training, data protection, monitoring, BCDR and policy enforcement.

Security standards, which is *a detailed, prescriptive document that specifies the technical or operational requirements for implementing specific security measures or achieving compliance with a policy*, must secure the working environment on the cloud without affecting its performance and reliability. Security policies work accordingly to regulatory authorities, covering various service-level agreements, issues in client management, and antecedent trust. Security policies and security standards are complementary, and standards can explicate how to comply to or satisfy a policy. Some security services and topics follow, which are connected to cloud security:

2.2.2.1 Service-level agreement (SLA)

An SLA is related to the customer-provider relationship, so it takes both of them into account and their concerns. It is expected that every service provider will use SLAs to manage customer expectations. In fact, the SLAs state the conditions under which providers will not be held liable for outages or performance issues. They also indicate the conditions under which providers will pay penalties as they will be liable for the occurring outage or performance issue. Also, SLAs are a representation of the performance characteristics of services that may be used for comparison purposes. SLAs could cover availability, which is related to security, or they could also include other security-related guarantees. When SLAs include such kind of clauses, they can be used as a cloud security policy tool.

2.2.2.2 Client management issues

[8] Ahmed M, Litchfield AT (2018), “*Taxonomy for identification of security issues in cloud computing environments*”, J Comput Inf Syst 58(1):79–88

The problems of client management as one of the most important concerns in cloud security have many different aspects, such as client experience (CX), client-centric privacy, client authentication, and client SLA. These aspects will be discussed in the paragraphs that follow.

(i) Client experience (CX). The customers expect the provider to satisfy special demands, personal conditions, as well as overcome the life difficulties of each customer. Providers can offer only better services if they know what their customers need and deliver customized solutions for them. Thus, customer experience in the cloud must be taken into consideration. This experience makes the companies able to deliver products and services according to customer values. With cyberattacks and privacy breaches becoming increasingly sophisticated, businesses must reinvent their security strategies to stay ahead since customers know the importance of securely handling their sensitive information. Organizations must also ensure they are not ignoring the importance of a frictionless user experience while adopting multi-layered security mechanisms. Hence, enterprises should blend customer experience with security to please their customers.

(ii) Client authentication. Ensure that users access a server or a remote computer by providing a digital certificate (or credentials); the digital certificate would act like a kind of "Digital ID." This will be used to cryptographically bind customers' and employees' identities. In return, the digital certificate would enable the cloud resources (i.e., cloud security/authentication/access control mechanisms) to manage access and prevent wrongful users from accessing the cloud services. Users should be able to get access to their purchased or leashed cloud services from anywhere and any device like cell phones, tablets, laptops, smart TVs, etc.; however, it becomes a challenge for cloud service providers to guarantee that only valid users get access to their services[9].

The authors in (Revar and Bhavsar, 2011) investigate existing authentication mechanisms, such as SSO. The authentication of users through mechanisms, such as cURL and SSL is very effective as it provides efficient solutions to protect users over insecure networks. However, the authors did not provide how the SSO can access cloud services from the mobile or other cloud compatible device, due to vulnerabilities of the cloud computing (shadow IT, lack of visibility).

The authentication scenario of SaaS based applications is shown on Fig. 3.

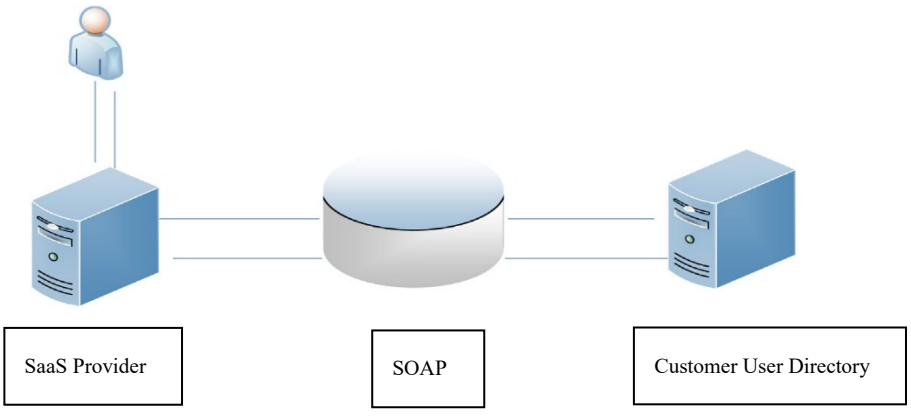

Fig. 3. Simple version of SSO

[9] Fernandes DA, Soares LF, Gomes JV, Freire MM, Inácio PR (2014), “*Security issues in cloud environments: a survey*”. Int J Inf Secur 13(2):113–170

- The user requires authenticating himself while accessing the information from SaaS provider. As the SaaS platform is web based therefore some sort of encryption is done in the URL or a cookie.
- Through a direct web service call, the information is then authenticated against the customer’s user directory (e.g., LDAP)
- The customer’s user directory replies back with some sort of authorization and authentication information.
- Finally, based on the scenario of authentication and authorization queries, the resulting request will be fulfilled or denied.

(iii) Client-centric privacy. Nowadays, the most significant cloud vulnerabilities appear to show up when privacy is lost, and data leakage is accomplished. In fact, data processing and storage by a provider causes this challenge and compromises client-centric privacy[10]. Implementing a customer-centric approach to data privacy involves more than just legal compliance. It requires businesses to prioritize the customer's right to control their data processing and actively engage in practices that promote transparency and trust. When the users submit their data into a cloud service to be processed, for instance, it has no control over the procedure of processing. This is in fact one major challenge in adapting cloud services (Subashini et al, 2011).

(iv) Service-level management. As mentioned, an SLA is used to provide the agreement between a cloud service provider and consumer as well as the cloud service cost. In fact, service-level management (SLM) is the process of “monitoring” the SLO assessment obligations and who must fulfil them for the level of service that is required. This process may also measure the performance compliance expected by the customer.

2.2.2.3 Antecedent trust

While setting up business relationships, a lot of facilitations are required, and one of them is the aspect of trust. Since pale attention has been received by cloud computing on the topic of trust, lack of trust understanding in cloud services has caused huge challenges in the creation of cloud computing adoption. Various aspects of security in cloud business, namely "Human factor", "Digital forensics", "Customer trust status", "Trust third party (TTP)", and "Governance" are referred in the following paragraphs.

(i) Human factors. The security reinforcement of the cloud, in fact, has given a secure infrastructure to information. Considering the human factor of cloud security is a must for an organization. In characterizing the human role in the security aspect of the cloud, one should note that people have been able to create various innovations and solve all the problems. *Social engineering* exploits human trust and is used for gaining initial access into an environment, collecting sensitive data, and/or performing malicious activities such as defaming or damaging a corporation’s reputation. As social engineering relates to cybersecurity, it is commonly defined as coercing/manipulating someone with the goal of obtaining valuable information. In essence, social

[10] Subashini S, Kavitha V (2011) “*A survey on security issues in service delivery models of cloud computing*”. J Netw Comput Appl 34(1):1–11

engineering is aimed at exploiting human trust. There are several tactics utilized by hackers to acquire the "valuable" information. These tactics might be used on their own or chained together.

(ii) Digital forensics. More network applications mean more digital crimes[11] (Safiia, 2015). Digital forensics play an important role in safeguarding and recovering data in operation. In other terms, forensic science in the digital world involves the procedure of searching and explaining electronic data. Digital forensics raises privacy concerns because it permits very detailed and thorough examinations of digital data. It can retrieve any digital artifacts left behind on a computer or device, including deleted data and fragments of data, not just those retrievable by the operating system or software, since it bypasses the usual access controls on computers. This means that data that users thought was hidden or long gone could be found in a forensic investigation, more so than with Web postings and electronic mail.

It thus may be true that with an increased number of users in the digital world, it becomes more potential that non-rightful malicious users could also exploit cloud services. Usually, cloud APIs are important tools for getting to and retrieving digital data in cloud forensics. With these APIs, forensic experts can connect directly to cloud platforms and get logs, information, and other important data without putting the cloud infrastructure at risk and violating data privacy legislation.

(iii) Costumer trust status. According to a report just published[12] (Spiceworks, "Diving into IT cloud services"), many organizations are concerned about their personal and sensitive data kept by companies. According to this report, while 93% of organizations use cloud services now, only 23% fully trust the public cloud to keep their data secure. The reason is quite simple: they are afraid that cloud-related companies (e.g., CSPs or cooperating third parties) can use the data for purposes other than those that they collect them for. When moving to IaaS, arguably the biggest wish from related authorities is having coherent security controls for integrated security, with centralized management across all cloud and traditional data center infrastructure.

(iv) Trust third parties (TTP). One of the key issues facing cloud users is how their data will be utilized because it is possible that malicious/curious cloud providers will use this information. A TTP in this regard can authenticate, audit, and authorize access to confidential data and protect information against unauthorized curious individuals (or programs)[13]. Therefore, the cloud environment will be threatened severely when the TTP is discredited, such that many security properties are compromised. Actually, the third-party entity is considered trusted but in essence this is at the surface and internally it might also act in a malicious or curious way (or could be penetrated such that issues like exposal of sensitive information can take place).

(iv) Governance. Cloud Computing governance is a branch of IT governance that introduces the use of integrated management with automated performance resolution, hence balancing resources in a cloud environment. Unfortunately, many organizations adopt a cloud environment without understanding the importance of governance in IaaS and the other cloud service types. If

[11] Safiia, M., "*An introduction to digital crimes*", International Journal in Foundations of Computer Science & Technology (IJFCST), Vol.5, No.3, May 2015

[12] Spiceworks, "*Diving into IT cloud services. More IT pros than ever are venturing into the cloud*", https://www.spiceworks.com/sw-marketing/reports/it-cloud-services/

[13] Basu S et al (2018) "*Cloud computing security challenges and solutions—a survey*". In: Proceedings of the IEEE 8th Annual on Computing and Communication Workshop and Conference (CCWC), pp 347–356

this issue is ignored, the administrator operations and security controls will be lost. Cloud governance has the best practices which are guidelines and strategies designed to effectively manage and optimize cloud resources, ensure security, and align cloud operations with business objectives. Cloud governance is also made up of cloud security controls, processes, and policies that help improve the management of cloud infrastructure.

Cloud security controls are deployed to safeguard an organization's data from distributed denial of service (DDoS) attacks, hackers, malware, and unauthorized user access while also supporting regulatory data compliance. In summary, cloud computing governance defines the policies and principles that regulate the life cycle of services offered in the cloud.

2.2.3 Data states

Data states include data ‘in transit’, in use’ and ‘at rest’. Data in transit is the data being transmitted from one node to another. This type of data travels internally between nodes within the same network or externally between nodes that belong to different networks. Data in use is the data that is accessible to the user (it can also be accessible by programs for processing reasons) in the form of documents, emails and applications. This type of data appears in plain text, so it can be easily interpreted and processed. Data at rest is the type of data that is stored in repositories. It consists of application databases, backup files and file systems. It is normally protected by strong access controls, including physical and logical mechanisms.

2.2.4 Cloud privacy

From the perspective of a common reader of the aforementioned NIST cloud computing definition, someone can extract several interesting keywords that highlight the great advantages of cloud computing like “convenient on-demand access” “configurable computing resources,” “minimal management effort” and “elasticity.” However, from the perspective of privacy, there is only one simple keyword in this definition that reveals one of the biggest concerns within the cloud computing concept, which is “*shared pool*”. Besides, a definition of privacy could be as follows: “*the right of a person to not have his/her private data brought to the public knowledge*” (Ghorbel et al., 2017).

In fact, the cloud computing paradigm offers several services for processing customer data on machines that the customer does not control and, even more, does not operate. Additionally, we have also to consider that private data might be selectively exposed to a specially crafted list of third parties and also controlled such that the processing and access of the person’s data is controlled by its owner, i.e., the person him/herself.

We mean by *private data* any *personal information*, habits, lifestyle, health data, etc. Private data, also known as proprietary or confidential data, refers to information that must not be publicly available and must be protected due to privacy, security, or legal concerns. This category includes personal information, such as social security numbers, financial records, medical history, and proprietary business data (when it concerns personal information). Also, '*Personal data*' *means* any information relating to an identified or identifiable natural person ('*data* subject') (GDPR, Art. 4(1)). This includes all data which is or can be assigned to a person in any kind of way. For example, the

telephone, credit card or personnel number of a person, account data, number plate, appearance, customer number or address are all personal data.

With the NIST cloud definition, a big challenge is raised: how can we be sure that private data put in a cloud shared pool will not be brought to public knowledge[14]?

Privacy, essentially, can be identified as the user's ability to maintain control over his/her data, and stopping it from falling into the wrong hands, may that be through a breach, leak, or cyber-attack, or by intentional improper actions by an entity that has the right to view and process the data. Security and privacy are generally mixed, but a security breach and a privacy breach may have different impacts. In any case, a privacy breach is always under the auspices of a security breach. Technically, there is a distinction between a security breach and a data breach. A security breach is effectively a break-in, whereas a data breach is defined as the cybercriminal getting away with information.

Cloud data security is the practice of protecting data and other digital information assets from security threats, human error, and insider threats. It leverages technology, policies, and processes to keep your data confidential and still accessible to those who need it in cloud-based environments.

Another aspect embodying *privacy issues* resides in the multi-tenancy feature of the cloud. It means that multiple customers can be served from the same instance of resource/service, by "securely" separating the resources on the logical/virtual level. This feature can achieve higher profit margin by maximizing resource usage and reducing cost through economies of scale. However, this gain cannot crop up without risking the leakage of confidential or sensitive data (e.g., due to improper resource isolation).

Confidential Data is generally non-public information, but not always inherently risky. It can include things like business plans or credit card numbers. *Sensitive Data* is highly confidential information that could cause significant harm if leaked.

'Personal information' is another important term that may be used in a slightly different manner by different people (Pearson, 2009), which needs to be protected and includes the following[15]:

- Personally identifiable information (PII), which is any information connected to a specific individual that can be used to uncover that individual's identity, such as their social security number, full name, email address or phone number, and
- Sensitive information, which is data that must be protected from unauthorized access to safeguard the privacy or security of an individual or organization.

Privacy is extremely important to every business and individual concerned about protecting confidential and personal information. Data privacy concerning *businesses* is the protection of personal data from those who should not have access to it and the ability of individuals to determine who can access their personal information (which is managed by business). Also, business secrets

[14] See also: Ghorbel, A., Ghorbel, M., Jmaiel, M., "*Privacy in cloud computing environments: a survey and research challenges",* Published online: 23 January 2017, Springer Science+Business Media New York 2017 Springer Science+Business Media New York 2017, DOI DOI 10.1007/s11227-016-1953-y

[15] Pearson, S., (2009, May), "*Taking account of privacy when designing cloud computing services*". In: Proceedings of the 2009 ICSE workshop on software engineering challenges of cloud computing. IEEE computer society, pp 44–52

can be considered as data which has to be protected for their privacy. Around the world, this has driven to the laying down of a large number of laws and legislations. Indeed, most of the national governments have imposed their local privacy legislation, but the existence of laws, even international laws (e.g., GDPR) is not always sufficient to protect private data/information

Security and privacy are two related but different terms. In fact, security represents a set of practices that are designed to ensure the confidentiality, availability and integrity of data (of any kind, either personal or non-personal) whereas privacy is defined as the appropriate handling (collection, use, share and storage) of personal data under specific circumstances (e.g., customer and/or law policies). Security techniques can be employed to ensure the appropriate use of data according to the required circumstances. Therefore, security is not privacy, but it is a basic foundation, among others, that helps to preserve privacy. See cloud security and related issues in Sections 2.2.1-2.2.2 of this chapter. However, apart from security measures, privacy-preserving techniques and measures can be employed to ensure/preserve privacy.

2.2.5 Cloud privacy issues and policies.

To benefit from advantages offered by CC, including high-level security, the user is led to move his private data to the CSP. Outsourcing this data to an external party provokes several problems and concerns such as: Who has access to the data? Where is the data stored? How many copies of data exist in the cloud? How to be sure that all copies are deleted when requested? Are the data involved persons (or entities) compliant with laws and user privacy policies in the cloud? All these issues are mainly raised due to: (i) the lack of user control, (ii) the dynamic nature of the cloud, (iii) the lack of technologies to ensure compliance with laws and user’s preferences and (iv) the difficulty in achieving accountability in the cloud environments. To better understand these cloud privacy issues, we will analyze them in the following paragraphs (of the next subsection).

The usage of the cloud covers several domains and levels; thus, several kinds of actors could have a relation with the life cycle of personal data in the cloud. Some possible actors that are involved in the data usage process in the cloud are the following:

Data owner. This is the main actor. She is the owner of private data and has chosen or not to host her data in the cloud or to use a data-related service (e.g., a data processing or storage one) that is hosted in the cloud. It should be noted that in some cases, the data owners are several individuals, especially when the data are co-produced by several individuals or in the case of data collection regarding several individuals (like usage databases, statistical data, etc.). As already mentioned before, the data owner has the right to control the processing and storage of his/her data. Such a right could be expressed in the form of privacy policies that need to be respected by the data processors.

Cloud service provider (CSP). The CSP is a key actor since it is the data hosting entity. It could be embodied by both individuals (internal employees) or/and software entities (web services interfaces, distribution, backup and load-balancing algorithms, etc.). The CSP can perform cloud-specific processing on private data, such as duplication, transfer, storage, research, indexation, statistic, mining, and segmentation. Obviously, it is possible that a CSP (as witnessed in reality)

could also perform undeclared actions that can threaten the privacy of data (e.g., sharing data with unauthorized third parties, sell or exploit enterprise secrets, etc.).

Cloud service broker (CSB). This actor is a company that adds value to existing cloud services (especially in public cloud) on behalf of customers of these services. The CSB uses three main roles including aggregation, integration and customization brokerage. In some limited cases, the role of the CSB is limited to a consulting role. It is possible that a cloud broker could also enhance or harm cloud privacy

Cloud-based service. This service represents applications and programs deployed in the cloud and which are used to accomplish a service, such as CRM (customer relationship management), document management, cloud storage service, etc. Since cloud privacy involves safeguarding sensitive data stored, processed, and managed in cloud environments by implementing effective cloud data protection practices, this actor contributes also to privacy.

Cloud-based service devops. A devops is a role in software development, who is responsible for the operation and improvement of a service. It maps to the owner of the service. It can be employed/played by the CSP or another party. Its role is to control and to improve the service. The devops of the service can have access to the service database and then can disclose private customer data.

Cloud-based service third party. They are actors implicated in the usage of the service deployed in the cloud. For example, an ecommerce SaaS application/service, a third party could provide electronic payment services. The third party that might provide assistance or services to a cloud service to facilitate its operation. Regarding personal data, these might be shared also between the cloud service provider and the third party. So, if the data owner is redirected to the respective service of the third party, these data are also shared by the data owner, the third-party and the CSP.

In the public cloud, private data are stored in remote machines controlled by the CSP. There is usually a lack of transparency of data processing, storage location, the number of data copies, etc. It is even difficult, and sometimes impossible, to know whether there were violations of privacy and who actor is responsible for them, as aforementioned.

The dynamic nature of cloud causes major problems for data privacy. Transborder data flow is one of these issues[16]. In fact, dynamic algorithms[17] are responsible for transfer and storage in the cloud, and thus, the user does not have any information about the location of his/her data. This issue is extremely relevant especially when data is handled in plain text form. In fact, data can be stored or processed in countries whose laws bring more privacy risks. Then, data became more susceptible to data disclosure issues.

Data replication is another problem that is caused by the dynamic nature of the cloud. Indeed, dynamic algorithms perform duplication of data in many servers to ensure data and service

[16] "Transborder data flow"("TDF") is a recently coined term that refers to the transmission of data or information over national. boundaries. Transmission can be affected in a variety of ways, including the telephone and the mails. Detecting credit card fraud at the point of sale offers one of the clearest examples of the benefits of cross-border data flows. No matter where you are in the world, your bank's computer back home can analyze your purchase and location in a matter of seconds when you swipe your credit card.

[17] Two algorithms are proposed to adapt the replication and placement of the data objects according to its popularity in the OSN. The first algorithm is Dynamic Fixed Time (DFT) which uses fixed time periods to adapt replication and placement. The second algorithm is Dynamic Exponential Time (DET) which determines the data object replication and placement based on exponential time periods.

availability. Hence, this joins the transborder data flow problem as it is difficult to guarantee that a copy of data or its backup will not be stored or processed in some countries forbidden by privacy policies or laws/legislation.

Another key concern in the public cloud is Privacy Compliance (PC), which indicates compliance with privacy law/legislation. This concern can cover the privacy issues as they can be regulated by expressing well-defined privacy policies (data owner preferences and/or regulations) and enforcing them.

To conclude, CC can magnify the existing privacy issues especially due to the lack of transparency concerning data handling and storage, the dynamic nature of the cloud, the lack of means for enforcement of privacy policies and the difficulty to achieve accountability in such an environment. Hence, it is logical that cloud customers' concerns about their sensitive outsourced data are more and more increasing.

2.2.6 Data loss (leakage)

Data leakage (or data loss) is a term used in the information security field to describe unwanted disclosure of information. This problem is mitigated by using different DLP methods and techniques, including both administrative and technical approaches. DLPs are defined as *the designated analytical systems used to protect data from unauthorized disclosure at all states using remedial actions triggered by a set of rules* (Alneyadi, S., et al, 2014[18]). This definition contains three main attributes that distinguish DLPs from conventional security measures. First, DLPs can analyze the content of confidential data and the surrounding context. Second, DLPs can be deployed to provide protection of confidential data in different states, that is, in transit, in use and at rest. The third attribute is the ability to protect data through various remedial actions, such as notifying, auditing, blocking, encrypting and quarantining. The protection normally starts with the ability to detect potential leaks through heuristics, rules, patterns and fingerprints. The prevention then happens accordingly.

2.2.7 Data requirements for cloud privacy

The security to the data leakage in the cloud can be addressed by using data confidentiality, data integrity, data availability, secure data sharing and privacy protection.

Data Confidentiality: Data confidentiality refers to protecting customer information or data, from computations by the cloud vendors and other users. Confidentiality is a major security concern in CC. Most of the cloud servers are managed and controlled by cloud providers who are not so trustworthy as being third parties (in literature, they are expressed as "curious actors"). As a result, confidentiality of customer data has many security concerns. Data confidentiality is important for users to store their private or confidential data in the cloud. So, data confidentiality, authentication, and access control issues in cloud computing must be addressed by increasing cloud reliability and trustworthiness.

[18] Alneyadi S., Sithirasenan E., Muthukkumarasamy V., "*A survey on data leakage prevention systems*", Journal of Network and Computer Applications, 2016, Elsevier, http://dx.doi.org/10.1016/j.jnca.2016.01.008

Data Integrity: Data integrity specifies the customer data or information in cloud server does not have any kind of unauthorized deletion, modification, or fabrication. Data integrity is one of the most critical elements in any information system. It is quite important that every organization, for its specific enterprise cloud resources to ensure that valuable data and services are not abused, misappropriated, or stolen.

A similar (although different) category to Data Integrity is Computational Integrity. Computational integrity implies that any of the programs executed are not compromised by malicious users. The attackers mostly target *untrusted servers* and take advantage of cloud owners' loss of control (due to shared responsibility principle and the multi-tenancy feature) which leads to security concerns. The lack of trust arises naturally when the database server is owned by a third party, as in the case of cloud computing. It also arises if the server may have been compromised, or there is a malicious insider. Ensuring the trustworthiness of data retrieved from such an untrusted database is of utmost importance. Therefore, attackers can attack such servers and then focus on violating the integrity of data and programs

Data Availability: Data availability refers to cloud services which are available all the time or not to cloud customers. Attackers use DoS attacks and make the cloud services unavailable to the user. Data availability means, when accidents, such as hard disk damage, IDC fire, and network failures, occur, the extent that the user's data can be accessed when requested or recovered and as well as the way users can verify their data (i.e., that all data parts are in place and nothing has been lost or stolen) (especially via specific techniques rather than depending on the credit guarantee by the cloud service provider alone).

Moreover, CSPs should ensure data security, particularly data confidentiality and integrity, as well as explain the jurisdiction of local laws to the clients. Further, depending on the type of service, the CSPs should share all such concerns with the client and build trust in this connection.

Data Sharing: Data sharing is becoming increasingly important for many users and sometimes a crucial requirement, especially for businesses and organizations aiming to gain profit. Historically, many people viewed the computer as “impersonal giants” who threatened to cut jobs of many people through automation. To enable data sharing in the Cloud, it is imperative that only authorized users are able to get access to data stored in the Cloud. The ideal requirements for data sharing in the Cloud are below:

• The data owner should be able to specify a group of users that are allowed to view his/her data.

• Any member of the group should gain access to the data anytime without the data owner's intervention.

• No other user, other than the data owner and the members of the group, should gain access to the data, including the Cloud Service Provider.

• The data owner should be able to revoke access to data for any member of the group.

• The data owner should be able to add members to the group.

• No member of the group should be allowed to revoke the rights of other members of the group or join new users to the group.

• The data owner should be able to specify who has read/write permissions on the data owner’s files

Data Accountability: One common definition of data accountability is given by the Galway Project[19] in the context of corporate data governance state that: “Accountability is the obligation to act as a responsible steward of the personal information of others, to take responsibility for the protection and appropriate use of that information beyond mere legal requirements, and to be accountable for any misuse of that information.” Hence, the way to achieve accountability in the computing context is to introduce a strong emphasis on auditing mechanisms.

Data Privacy: Data Privacy refers to the protection of customer data which is stored at the cloud provider’s end. Due to untrustful servers and insider attackers at cloud providers end, security concerns related to privacy exist in CC.

Cloud data privacy is all about safeguarding any data in the cloud from loss, leakage, or misuse through breaches, exfiltration, and unauthorized access. As more organizations shift their workloads to the cloud, addressing cloud data security concerns becomes increasingly essential. Therefore, Data privacy is the protection of personal data from those who should not have access to it and the ability of individuals to determine who can access their personal information.

[19] It is an effort initiated in January 2009 by an international group of experts from government, industry and academia to define the essential elements of accountability and consider how an accountability approach to information privacy protection would work in practice

3 *Methodology review*

The goal of this chapter is to unveil the review methodology followed, i.e., the process that was used to select articles in the context of answering research questions that relate to the main objective of this thesis. In addition, this chapter explains in which places of the remaining report are the questions answered while it supplies some initial quantitative analysis results.

3.1 Literature sources, research questions to answer and methodology

In this respect, the essential methodology—the theoretical framework for this academic paper — was developed with three distinct and subsequent cycles of extensive literature review, which in turn is informed by existing literature from notable scholars, such as Webster & Watson and Mueller-Bloch & Kranz, as well as, Google Scholar, Web of Science plus particular search mechanisms of well-known editors like Springer, ACM and IEEE.

As a prerequisite step before coming up with our research questions to be explored from the literature sources, we went ahead to conduct an advanced preliminary research phase to understand the major topics. This was an important first step, as this helped us establish a starting point for our understanding of the subjects under our study. In doing so, we found some knowledge gaps and limitations identified in our selected topics through a preliminary literature preview that further highlighted those areas that needed to be explored further. Upon attaining a solid understanding on the broader themes we were interested in, we then narrowed our focus to explore a specific scope of study that would enable us to go deeper into the exploration of the topics we had chosen. The literature results from our methodology on the topics (risks, attacks, vulnerabilities, countermeasures) in relation to the following questions, led us to the next step of our research: to design a set of filtering criteria (see Section 3.3). We finally took time to review our questions (see Table 2) and ensured that they were clear, focused, and specific, offering a deep level of scientific challenge. The framed research questions managed to incorporate both the broad and narrowly focused research aspects.

The research questions, positioned in specific categories and related to the main objectives of the literature review, we asked about the literature sources to clarify the most proper and relative patterns of attacks, vulnerabilities and countermeasures for cloud privacy, were as follows (Table 1):

Table 1. Type of specific research questions (answered in specific chapters in the thesis)

Descriptive research questions	What are the main attacks and threats regarding privacy in the cloud? What are the current research gaps in preserving cloud privacy? What are the main research directions to follow to cover these gaps?	Ch. 4 – Ch. 6
Correlational research question	Is there any countermeasure for protecting privacy in the cloud for every type of attack and threat?	Ch. 5
Comparative/Exploratory research question	Are there one or more countermeasures protecting cloud privacy that can be considered better than the other?	Ch. 5
Evaluation research question	What is the level of effectiveness of countermeasures in terms of cloud privacy?	Ch. 5
Explanatory research question	What is the reason for the inefficiency and ineffectiveness of countermeasures in protecting privacy in the cloud for which future trends have been developed?	Ch. 6

The proposed literature review method is based on iterative cycles, in three steps: Initially, findings (based on research publications) from key words (for threats, attacks, and countermeasures) are selected, then based on the results for every concept, in the second step, a selection of literature related to the analysis of privacy issues and attacks is considered (filtering mappings) and finally a selection and synthesis of the previous step findings are considered in a matrix, based on the threats, attacks and the proposed countermeasures and their classification. Under this method, we will make a synthesis towards a proposed categorization of countermeasures related to the main research goal of using the proper countermeasure to the most common attacks or threats in cloud privacy.

All these iterative cycles, towards the thesis’ specific objective, are crucial in providing knowledge contributions that give enough insights on topics dealing with threats, attacks, and countermeasures within a cloud computing context. Nevertheless, the first cycle focuses on the "threats, attacks, and countermeasures" areas related to the cloud sphere. It becomes clear from the outset that several literature review studies are needed for critical evaluation and analysis in-depth to draw relevant insights for developing this research. By conducting such an in-depth examination, core concepts that appear to emerge from all sources reviewed are brought out using the rigorous methodology proposed by Vom Brocke et al (2009)[20]. Vom Brocke et al. (2009) framework suggests that the process of creating a research agenda starts after the literature synthesis has been undertaken. The chart method can be used for literature synthesis and for research gap localization (Vom Brocke et al. 2009; Webster and Watson 2002). When researchers start to chart the reviewed literature according to various concepts, they already perceive “blank fields” (no results) in the chart, which may indicate research gaps. The result of this first literature review is a *concept matrix*, which structures and supports the subsequent literature review cycles (Fig. 4).

Proceeding to the second cycle of literature review based on the results from the first step, we identify key problem areas: *threats, attacks and countermeasures in cloud*. Based on the identified concepts that constitute the aforementioned concept matrix, multiple key word searches

[20] Brocke, J.V., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., Cleven, A.: “*Reconstructing the giant: on the importance of rigour in documenting the literature search process*”. Presented at the http://www.alexandria.unisg.ch/Publikationen/67910 June 10 (2009)

are performed based on the concepts. Based, also, on literature findings for each concept, we classified the threats (building an own model), and we filtered the results (for attacks, threats and vulnerabilities) in sub-sections 3.2 and 3.3 (in Ch. 4, we fully analyze them). The resulting artifact is the relevant sources of literature and their corresponding literature concepts for the key words and privacy issues areas. The outcome of the search part of the methodology is a set of articles related to the key issues, general countermeasures and the relevant problems (*problems matrix*). We also present a similar table (Table 2) with relevant articles regarding cloud service models, taxonomy for algorithms in cloud computing, types of attacks and their categorization, etc. We used a similar concept to this table, to construct our initial concept matrix and a problems matrix.

Table 2. Existing surveys on cloud service models (Iqbal, S. et al., 2016)

Papers	Description	Parameters	Year
(Juliadotter and Choo, 2015)	The authors provide a comprehensive study on taxonomies, and present different attacks against cloud services.	Attack taxonomy classifiers	2016
(Shameli-Sendi et al., 2015)	This research studies DDoS against CC as well as the mitigation strategies.	DDoS mitigation	2015
(Alizadeh et al., 2015)	The research defines the authentication methods and compares the mobile CC authentication methods by considering five evaluation metrics.	Cloud-side authentication methods, User-side authentication methods,	2015
(Shiraz et al., 2015)	The research investigates existing computational offloading methods and highlights technical issues such as resources intensive distributed platform, deployment of virtual mobile devices on cloud server node, accessibility of virtual mobile devices, and constraints on computing resources of the virtual mobile device.	Energy consumption and timing cost. Data transmission in computational offloading	2015
(Shaikh and Sasikumar, 2015)	This research presents the measurement to calculate the trust model. This trust model computes the trust value to strengthen the security.	Static and dynamic trust	2015
(Liu et al., 2015)	The authors present a thematic taxonomy for application partitioning algorithms in cloud computing.	Dynamic computational offloading	2015
(Ab Rahman and Choo, 2015)	The research surveys existing security incident handling and digital forensic in CC.	Digital forensic, incident handling, and Capability Maturity Model	2105
(Simou et al., 2014)	The paper presents CC forensic challenges.	Cloud Forensics, Cloud Forensics Process, Cloud Forensics Challenges, Digital Forensics	2014
(Toosi et al., 2014)	The survey paper discusses and categorizes possible cloud interoperability scenarios and architectures.	Cloud federation, Multi-cloud, Utility computing, Inter-cloud, cross-clouds	2014
(Whaiduzzaman et al., 2014)	The paper surveys vehicular CC issues.	Cloud formations, Key management, Inter cloud communication systems,	2014
(Manvi and Shyam, 2014)	This paper surveys resource management techniques in IaaS.	Resource provisioning, resource adaptation, resource allocation and resource mapping.	2014
(Patel et al., 2013)	The paper surveys and presents a conceptual Intrusion Detection and Prevention (IDP) architecture for CC.	Autonomic computing, risk management, ontology, and fuzzy theory	2013
(Grispos et al., 2013)	The paper presents the case study of Global Fortune 500 organizations and identifies the real world information security documentation issues.	Auditing policies, Standards and guidelines Applicable	2013
(Modi et al., 2013)	The paper surveys CC security issues.	Virtualization, Security, Privacy and Vulnerabilities	2013
(Sun et al., 2011)	The authors survey CC security, privacy and trust issues.	Presents the solution to analyze and eliminate potential privacy, security and trust threats.	2011
(Subashini and Kavitha, 2011)	The paper classifies CC security issues, focusing on Software as a Service.	SaaS	2011

The final literature reviews cycle maps potential solutions to the previously identified problem areas, namely with a set of countermeasures regarding privacy in cloud. These are specific countermeasures for cloud privacy while the countermeasures identified in previous cycles concern cloud security in general. In addition to rigorous sources from literature, some sources from EU reports (technical reports) as well as corresponding projects, were considered. The result is an extension of the problem matrix by mapping different solution components to each issue area, leading to a *solutions matrix*. Altogether the three cycles of literature review provide a set of concrete solutions for the vulnerabilities, attacks and impacts/risks, as long as countermeasures in cloud, supported by rigorous and relevant literature findings, for presenting a survey of methods protecting the privacy in cloud.

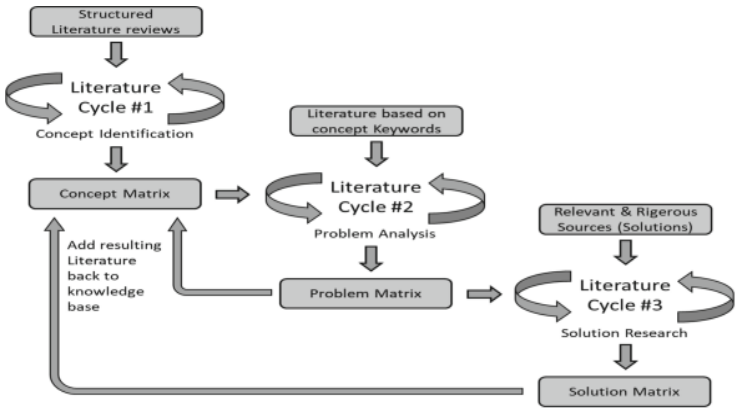

Fig. 4 Structure of the described literature review cycles

3.2 Classification models review of threats

Mostly, security attacks occur due to the behavior of legitimate users[21] (Iqbal et. al. 2016). These can, however, be vastly minimized and reduced by the thoughtfully intelligent inclusion of multiple security checks and measures. Such proactive steps go a long way in addressing the potential vulnerabilities and guarantee a much safer environment for one and all. We present some classification and our proposition for cloud security threats classification. In Ch. 4, we present the analysis of cloud privacy and security threats, based on our model.

3.2.1 Security threats classification

3.2.1.1 Threats classification methods

(I) The three orthogonal dimensional model

Researchers (Lukas Ruf et al., 2014) proposed a new threat model to categorize security cloud threats to improve the understanding of threats and improve the existing threat classification models. That works addresses this problem by introducing a three-dimensional model that subdivides threat space into subspaces according to three orthogonal dimensions labeled motivation, localization and agent:

a) *Threat agent* is an actor that imposes the threat on a specific asset of the system, which is represented by three classes: human, technological, and force majeure.

b) *Threat motivation* represents the cause of the creation of the threat; it is organized into two classes: deliberate and accidental threat.

c) *Threat localization* represents the origin of threats, either internal or external.

(II) Hybrid model for threat classification

[21] Salman Iqbal, Miss Laiha Mat Kiah, Babak Dhaghighi, Muzammil Hussain, Suleman Khan, Muhammad Khurram Khan, Kim-Kwang Raymond Choo, “*On cloud security attacks: A taxonomy and intrusion detection and prevention as a service”,* Journal of Network and Computer Applications 74(2016)98–120, Elsevier, p. 102, par. 3.1

Researchers (Sandro et al., 2014) proposed a hybrid model for information system and security cloud threat classification named the “information system security threat cube classification model” or C3 model. They consider three main criteria:

a) Security threat frequency: It depicts the frequency of security threat occurrence.

b) Area of security threat activity: It represents the domain that is being affected by threats like physical security, personnel security, communication and data security, and operational security.

c) Security threat source: It gives types of the threat’s source (ie, Destruction of information, Corruption of information, Theft or Loss of Information, Disclosure of Information).

(III) Information Security Threats Classification Pyramid model

Researchers (Mohammed Alhabeeb et al., 2014) present, a classification method for deliberate security threats in a hybrid model that is named “Information Security Threats Classification Pyramid”. It classifies deliberate threats based on three factors:

a) Attackers’ prior knowledge about the system: It represents how much the attacker knows about the system in terms of system hardware, software, employees and users’ knowledge.

b) Criticality of the area: It represents the criticality of parts of the system which might be affected by the threat.

c) Loss: It represents all losses that can occur in the system or to the organization (privacy, integrity, etc) due to the occurrence of the threat.

3.2.1.2 Classification models for threat impact

(I) STRIDE Model

Microsoft developed a classification method, called STRIDE, which is applied on the network, host, and application levels. STRIDE allows characterizing known threats according to the goals and purposes of the attacks (or motivation of the attacker). The STRIDE acronym is formed from the first letter of each of the following categories: Spoofing identity, Tampering with data, Repudiation, Information disclosure, Denial of service and Elevation of privilege. It is a goal-based approach, where an attempt is made to get inside the mind of the attacker. Moreover, the STRIDE Threat Modeling framework is a systematic approach used to identify and analyze potential security threats and vulnerabilities in software systems (Khan R. et al., 201722). STRIDE Model is also used for Cyber-Physical Systems (CPS) which integrate digital and physical components, such as smart cars, automated factories, and medical equipment. When compromised, these systems can cause real-world harm.

(II) ISO model

The ISO standard (ISO 7498-2) has listed five major security threat impacts and services as a reference model: Destruction of information and/or other resources, corruption or modification of

[22] R. Khan, K. McLaughlin, D. Laverty and S. Sezer, "*STRIDE-based threat modeling for cyber-physical systems*" 2017 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe), Turin, Italy, 2017, pp. 1-6, doi: 10.1109/ISGTEurope.2017.8260283.

information, theft, removal or loss of information and/or other resources, disclosure of information, and interruption of services (NECS-European Network on Cybersecurity, 2017[23]).

Most classifications of cloud security (and privacy) threats (like those previously analyzed) are usually limited to the use of one or two criteria to classify threats while the others present a non-exhaustive list of threat categories (not all threats are covered on classification) and their categories are not mutually exclusive. If a threat can be classified in multiple categories, this can cover in a better way dynamicity. On the other hand, by having very static mappings, it is rather strange and impossible to change the sole classification of a threat in a specific category. Thus, by having multiple mappings, it is tolerable to have one mapping removed or introducing a new mapping. This does not require changing the definition of a category. In static, mutually exclusive cases, if a threat is changed, it is possible that no classification could be performed without changing the definition of classes, which can be a major issue. In fact, organizations are prone to several kinds of cloud threats which affect their reputation. So, it is important that we can identify all threats characteristics to mitigate their risks.

3.3 Filtering criteria for results

In our thesis, we used specific criteria to filter the research results for all kinds of categories (attacks, threats, vulnerabilities) affecting cloud computing security, after the results from our methodology review were produced (presented previously). The criteria are categorized into *inclusion/exclusion* and *quality criteria*. These criteria include the following:

- *Inclusion/Exclusion:* we include articles that are highly related to the main objective and research questions of the literature review. Based on the three stages of the review methodology, we used for filtering in the literature databases, the *publication type* criterion (a publication type indicates the medium in which a specific work was published and maps to journal articles, conference articles, books, book chapters and thesis) while putting specific keywords regarding the subject of the thesis. Then, we kept the most relevant results to cloud computing and privacy issues. We also used the following inclusion criteria for the searched literature:

1. It must have been published in a certain timeframe (2012 and onwards: due to the relevance of technology advancements, public cloud" was in the air and the rush to infrastructure in the cloud)
2. From research libraries of International Organizations, European Union and from the scientific databases Elsevier, Springer, IEEE, Sciedu, EBSCO, Emerald
3. Apart from the aforementioned publication types, we also relied on relevant web pages.

By considering the last criterion, referring to relevant web pages is a helpful literature research tactic, to find real cases of attacks, threats and vulnerabilities, which might not be found in scientific journals. These cases are also important for the thesis as examples in their presentation regarding

[23] European Network on Cyber Security, NECS, Ref. Ares(2017)4315821 - 04/09/2017, p.22

attacks (see section 4.2.6). Moreover, relevant web pages include a real-world description of cloud privacy attacks, which are aimed by our research and enable the production of a more realistic presentation. At the same time, the ease of the description of complex technical details is also important. Finally, such web pages usually help to describe the tools and solutions which can be used to defeat the attacks.

Finally, a last exclusion criterion was that we exclude articles or web pages that are published in a language other than English and Greek.

- *Quality*: Regarding the related articles, after applying the aforementioned filters, we filtered the remaining articles based on their content. So, we excluded those articles that suggest countermeasures for cloud attacks which were not very well presented or were quite limited in terms of the categorization of the attacks/threats/vulnerabilities that they suggested or were not at all understandable. Moreover, we excluded articles that did not compare certain cases, be experimental or observational or both, use only qualitative methodology (we want quantitative data also).

3.4 Preliminary quantitative results

At the end, from the initially identified circa 200 sources by applying the three stages of the methodology, we examined and retained about 70 of them (35%) as the most related and qualitative research sources (quite recently published). From circa 70 selected sources/articles, 50 were scientific articles in journals (72%), 10 in conferences (14%), while the rest were technical reports (10 in number – 14%) (Fig. 5, b). The majority (72%) of the literature was mainly from Elsevier and other publishers (Fig. 5, a).

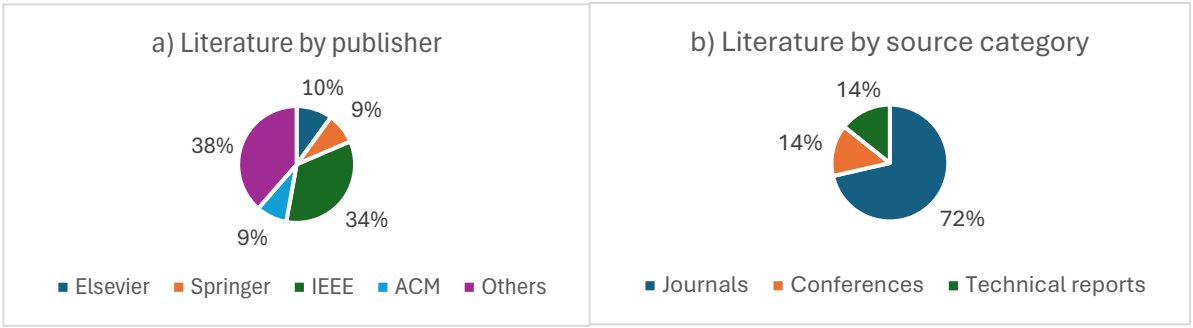

Fig. 5 Charts depicting the selected articles’ distribution

The majority (64%) of the selected (70) sources spanned over the period 2016 - 2022 (Fig. 6), as the trend line implies. Indeed, from 2016 to 2020 there is an increase in publications (of the related articles) due to technological improvements with a peak in 2020, probably due to the recent health crisis, during which major technological advances were taken place to meet protection and communication needs and to the plenty of attacks that happened at that period. Afterwards, a logical decrease in scientific publications up to nowadays, took place.

The high level of publications in cloud computing on 2012, is explained because, on this year the “public cloud” was in the air and the rush to infrastructure in the cloud happened on 2012, which was the biggest thing to happen to the cloud: the arrival of three major new IaaS (infrastructure as a service) players: Google Compute Engine, HP Cloud, and Microsoft Windows Azure.

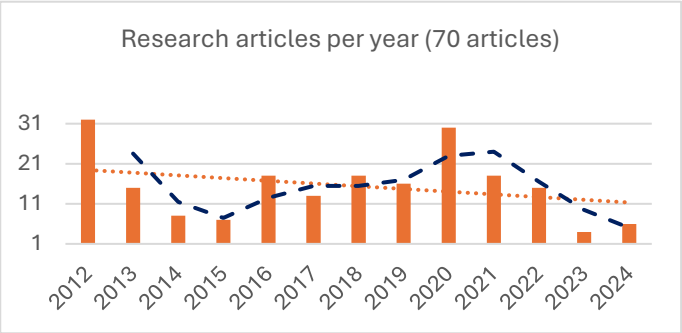

Fig. 6 Chart depicting articles’ statistics

4 *Privacy vulnerabilities, attacks, and threats in the Cloud*

In this chapter, we classify and analyze the attacks, threats and vulnerabilities in the Cloud regarding cloud privacy and cloud security. In particular, the cloud surface attacks are initially categorized as well as the attacks, threats and vulnerabilities in every cloud layer. Then, we present a reference to the major security attacks that affect cloud privacy, data leakage and data protection.

It must be noted that the presented taxonomy of vulnerabilities, attacks and threats in cloud privacy is necessary for the presentation of the proper countermeasures and their classification in the next chapter.

4.1 Classification and analysis of attack surfaces on Cloud Computing security

It is in this context that there is a dire need to develop and establish new taxonomies and classifications that help categorically understand the threats in question. In this paper, we present the *proposed cloud security/privacy threats model* as a taxonomy based on the CC service delivery model (Fig. 8). Cloud Security Alliance and Gartner[24] have identified several potential threats that may be countered in a CC environment. Security in CC environments needs a holistic approach. As there has been a paradigm shift to CC, new solutions are needed to support modern business functionality. However, this adaptation is not hazard free and poses new security threats to its adopters. Therefore, in Fig. 7 a proposed taxonomy of attacks along with their cloud service delivery model that they concern, is depicted.

[24] Council, S.-a.-a.-S.E., 2006. "*Software-as-a-Service; A Comprehensive Look at the Total Cost of Ownership of Software Applications*."

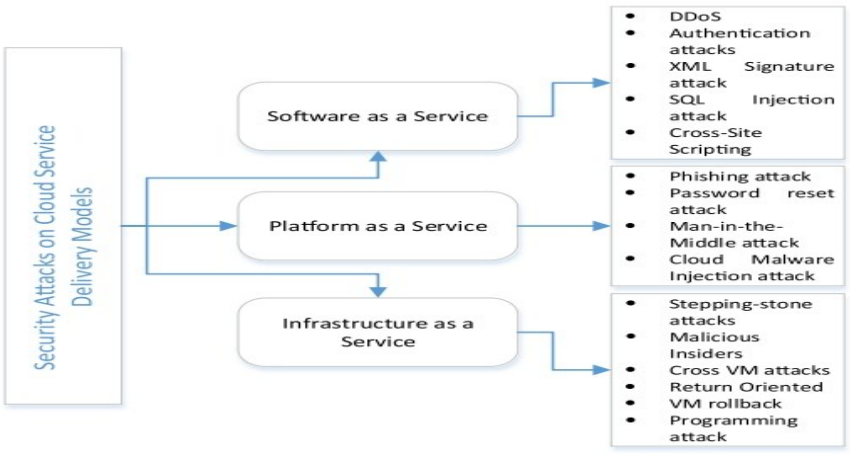

Fig. 7 Taxonomy of attacks along with their cloud service delivery models (Iqbal S. et. al,, 2016)

Proposed cloud security/privacy threats model

In this respect, we have come up with our own classification model (see Fig. 8) which includes the following criteria or dimensions:

- *Security threat source*: The origin of threat, which can be either internal or external. A cloud threat can be caused by internal, external or both external and internal entities. In this thesis, we focus only on a binary classification of the threat’s origin: internal or external, to localize the origin (or source) of a threat.

  (i) *Internal threats*. Internal threats occur when a respective entity has authorized digital access to the system with an account (e.g., on a specific server) or physical access to that system. A threat can be internal to the organization as the result of employee action or failure of an organization process. Internal threats have more rights in the system, so they can cause more damage to it. Thus, internal threats can have more severe impact than external threats.

  (ii) *External threats*. External threats can arise from individuals or organizations working outside of a company. Such entities do not have authorized access to the cloud environment or a very limited one. The most obvious external threats to (cloud) systems and resident data are natural disasters: i.e., hurricanes, and earthquakes. External attacks occur through connected networks (wired and wireless), physical intrusion, or a partner network.

- *Security threat agents*: These are the actors that impose a threat to the system. They can be classified into human, environmental and technological agents. The proposed classification covers the full set of potential agents, including humans, and human-made objects (technological, i.e. bots, malware, etc), and, natural for all those agents on which humans do not have any influence.

  (i) *Human threat agents*. This class includes a human (e.g., an employee or a hacker) as the main actor posing a threat through his/her potential actions over the system, which can be accidental or on-purpose. For instance, a recently fired employee might

take revenge by physically destroying part of a specific company facility, like some machines in a computer room (on which he/she has still access).

(ii) *Environmental threat agents*. The environment itself can be the source of a threat. Indeed, this class includes other threats, such as riots, wars, and terrorist attacks, which are essentially caused by humans, but occur as a threat in the external environment, as well as other threats like natural disasters (floods).

(iii) *Technological threat agent*. Technological threats are caused by physical and chemical processes on material (like a botnet which is a network of infected computers that work together to carry out an attacker's goals). It also includes indirect system support equipment like power supplies and issues like improper hw/sw configurations, hardware flaws, and power outages.

- *Security threat motivation and intension*: The goal of attackers on a system which can be malicious or non-malicious. Attackers normally have a specific goal or motive for an attack on a system. These goals can map to malicious or non-malicious actions and their results.

  (i) *Malicious threats* map to internal or external attacks conducted on purpose by human or software agents (e.g., viruses, Trojan horses, or worms) in order to harm and disrupt an organization.

  (ii) *Non-malicious threats* can occur due to, e.g., poor security policies and controls that allow vulnerabilities and errors to take place or human mistakes. They might be caused by ignorant employees with the aim of not harming the system or by the environment itself.

  The intent of the human who caused the threat that is intentional or accidental. This criterion allows us to reconstruct attack behaviors and full malicious behavior to understand its intention. It presents a predictable factor to help investigate a case with high accuracy and hence reduce risks and help to accelerate decision making for catching real agents. It includes Intentional Threats, that are result of a harmful decision and Unintentional Threat, which represents threats that are introduced without awareness.

- *Threats impact*: Threat impact is the main outcome that results from a successful attack that realizes a certain threat. For our model, we identified the following threat impacts (ISO enriched): *Destruction of information*, *Corruption of information*, *Theft/loss of information*, *Disclosure of information*, *Denial of service*, *Elevation of privilege* and *Illegal usage*.

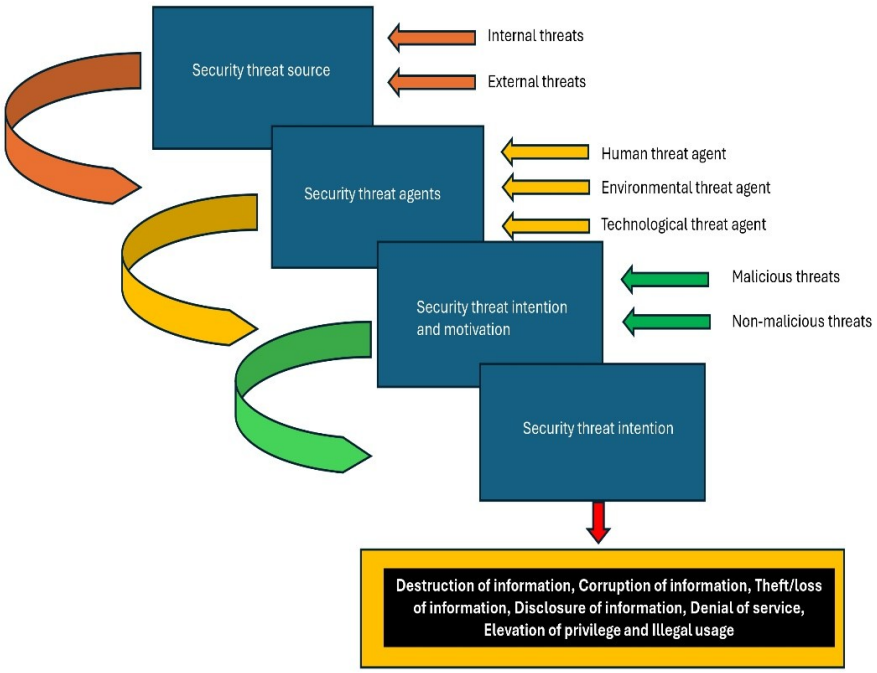

Fig.8 Classification model for cloud security/privacy threats

The aforementioned cloud threat classification model allows us to well define and articulate the main threat characteristics. Indeed, it serves as a guideline to determine what kind of cloud threats influence a system (like a cloud environment) while it also assists in the selection of security and privacy decisions and countermeasures. It also assists to realize threats’ techniques and their potential impacts, which in turn will be combined with our classification countermeasures’ model and to a proposed algorithm for the selection of the proper countermeasures (based on criteria convergence).

4.1.1 Cloud attack surface and security vulnerabilities’ classification

Cloud computing has revolutionized the way businesses and individuals store, access, and manage data. By offering on-demand computing services over the internet, cloud platforms provide scalability, cost-efficiency, and accessibility. However, despite these advantages, cloud environments are susceptible to a range of security vulnerabilities. These vulnerabilities arise due to shared infrastructure, remote accessibility, and the complexity of cloud configurations. Understanding these risks is crucial to implementing effective security measures and protecting sensitive data from cyber threats. We classify this kind of vulnerabilities (Fig. 9) into three categories, regarding the cloud layers[25], as follows:

[25] We classified the respective vulnerabilities for cloud security, taking under consideration our proposed model in Ch. 4, section 4.1

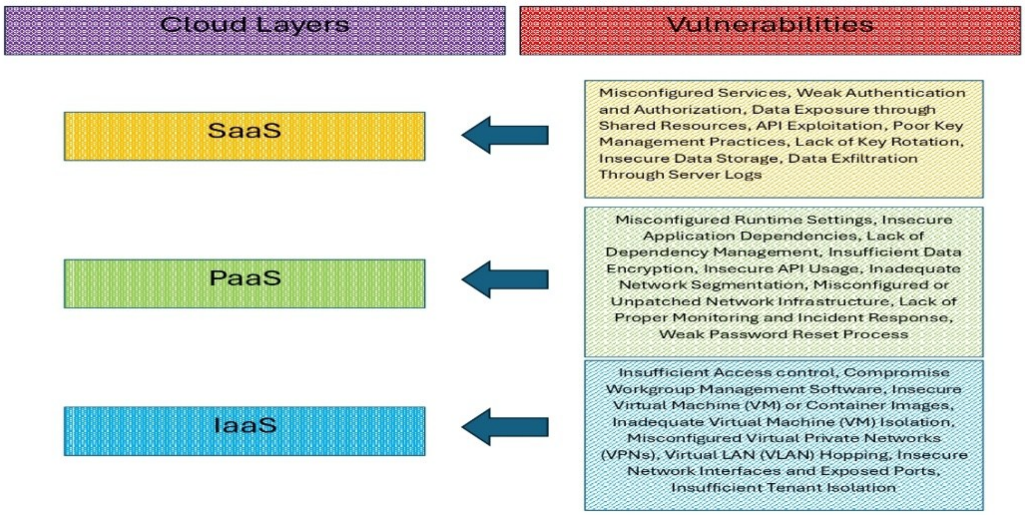

Fig 9. Security vulnerabilities per Cloud Layer

4.1.1.1 Vulnerabilities of the SaaS cloud layer

Software-as-a-Service (SaaS) provides cloud-hosted applications (e.g., Google Workspace, Microsoft 365, Salesforce), but it introduces security vulnerabilities due to misconfigurations, weak authentication, API security flaws, insider threats[26], and data exposure. Some of the most common vulnerabilities of this category are:

- *Misconfigured Services.* SaaS applications often come with customizable settings and configurations. Misconfigurations, such as overly permissive access controls or improperly set sharing permissions, can expose sensitive data or functionality.
- *Weak Authentication and Authorization:* Weak authentication mechanisms, such as simple passwords or lack of multi-factor authentication (MFA), make SaaS applications vulnerable to credential-based attacks. Poor authorization practices can also lead to privilege escalation.
- *Data Exposure through Shared Resources***:** SaaS platforms often rely on shared resources (e.g., databases, storage) to serve multiple tenants. Inadequate isolation between tenants can lead to accidental or intentional data exposure.
- *API Exploitation:* SaaS applications rely heavily on APIs for integration and functionality. Insecure APIs, such as those lacking proper authentication, encryption, or input validation, can be exploited by attackers.

[26] Insider threats are another serious concern, occurring when employees, contractors, or business partners misuse their access to harm the organization. This could include sabotaging systems, leaking proprietary business information, or deliberately disrupting operations. Insiders can also cause damage *unintentionally* by clicking on phishing emails, misconfiguring security settings, or mishandling sensitive data. Organizations must implement strict access controls and employee training to mitigate these risks.

- *Poor Key Management Practices:* SaaS applications often use encryption keys to protect data. Poor key management practices, such as storing keys in plaintext or using weak encryption algorithms, can undermine data security.
- *Lack of Key Rotation:* Failing to regularly rotate encryption keys increases the risk of key compromise. Long-lived keys are more likely to be exposed or cracked over time.
- *Insecure Data Storage:* SaaS applications may store data insecurely, such as using weak encryption or failing to encrypt sensitive data at rest. This makes data vulnerable to theft or unauthorized access.
- *Data Exfiltration Through Server Logs:* Server logs often contain sensitive information, such as user credentials, API keys, or personal data. If logs are not properly secured or monitored, attackers can exploit them to exfiltrate data.

4.1.1.2 Vulnerabilities of the PaaS cloud layer

PaaS enables developers to build, deploy, and manage applications without managing the underlying infrastructure. However, this abstraction of infrastructure management introduces unique security challenges. Below are some of the most common of the key security vulnerabilities associated with the PaaS cloud layer:

- *Misconfigured Runtime Settings:* PaaS platforms allow developers to configure runtime environments for their applications. Misconfigurations, such as overly permissive permissions or insecure default settings, can expose applications to attacks.
- *Insecure Application Dependencies:* PaaS applications often rely on third-party libraries, frameworks, and components. If these dependencies contain vulnerabilities, they can compromise the entire application.
- *Lack of Dependency Management:* Failing to regularly update and patch application dependencies can leave PaaS environments vulnerable to known exploits.
- *Insufficient Data Encryption:* Data stored or transmitted within a PaaS environment may not be adequately encrypted, leaving it vulnerable to interception or theft.
- *Insecure API Usage:* PaaS platforms rely heavily on APIs for integration and functionality consumption. Insecure APIs, such as those lacking proper authentication or input validation, can be exploited by attackers.
- *Inadequate Network Segmentation:* PaaS environments often host multiple applications or services. Inadequate network segmentation can allow attackers to move laterally within the environment.
- *Misconfigured or Unpatched Network Infrastructure:* PaaS providers manage the underlying network infrastructure, but misconfigurations or unpatched vulnerabilities can still pose risks. This depends on whether the PaaS provider owns the network infrastructure. If the infrastructure is supplied by an IaaS provider, the PaaS provider cannot manage the infrastructure. He/she can just slightly configure it.
- *Lack of Proper Monitoring and Incident Response:* Without robust monitoring and incident response capabilities, security incidents in PaaS environments may go undetected or unaddressed.

- *Weak Password Reset Process:* The password reset process may rely on weak authentication mechanisms, such as security questions or single-factor verification. Also, password reset functionality in PaaS platform (which offers APIs via which, in turn, users can interact) may generate insecure links, such as those with predictable tokens or insufficient expiration times So, if the user forgets his/her password to access such APIs, then this vulnerability can take place (*Insecure Password Reset Link)*.

4.1.1.3 Vulnerabilities of the IaaS cloud layer

Since IaaS is the CC model that provides virtualized computing resources over the internet, which offers scalability, flexibility, and cost-efficiency, it also introduces unique security vulnerabilities that organizations must address to protect their data, applications, and infrastructure. Below are some of the most common of the key security vulnerabilities associated with the IaaS cloud layer:

- *Insufficient Access Controls:* Weak access control mechanisms, such as overly permissive user roles or lack of role-based access control (RBAC), can allow unauthorized users to access critical resources.
- *Compromised Workgroup Management Software:* Workgroup management tools used to manage IaaS resources (e.g., orchestration tools, configuration management systems) can be targeted by attackers, if not properly secured.
- *Insecure Virtual Machine (VM) or Container Images:* VM or container images used to deploy workloads may contain vulnerabilities, misconfigurations, or malicious code.
- *Inadequate Virtual Machine (VM) Isolation:* Poor isolation between VMs on the same host can allow attackers to break out of a VM and access other VMs or the underlying hypervisor (it can also take place in the case of containers).
- *Misconfigured Virtual Private Networks (VPNs):* VPNs used to secure communication between IaaS resources may be misconfigured, such as using weak encryption or allowing unauthorized access.
- *Virtual LAN (VLAN) Hopping:* VLAN hopping occurs when an attacker exploits misconfigurations in network switches to gain access to restricted VLANs.
- *Insecure Network Interfaces and Exposed Ports:* Network interfaces and ports that are unnecessarily exposed or improperly secured can be targeted by attackers.
- *Insufficient Tenant Isolation:* IaaS environments often host multiple tenants on shared infrastructure. Inadequate isolation between tenants can lead to data leakage or unauthorized access. Tenant Isolation concerns logical separation between customers in multi-tenant environments. The other term “Inadequate VM Isolation” is about security between virtual machines on a shared physical infrastructure. Both require strict security measures, such as strong IAM policies, hypervisor security, and encryption, to prevent data leakage and unauthorized access. In other words, both are Weak Resource Isolation.

4.1.2 Attack vectors for cloud security and CC surface attacks

An attack surface encompasses all the (vulnerable) points in a specific system that an adversary or unauthorized user can exploit to gain access to that system and damage it. Resource sharing, in a cloud multi-tenant environment, is one of the most critical issues, which creates new attack vectors. An attack surface is all the points of entry and vulnerabilities an attacker can exploit to infiltrate a network or a system. Understanding a network's attack surface is critical — by knowing where the vulnerabilities are and monitoring it accordingly, an organization can reduce their attack surface and make it much harder for attackers to penetrate and compromise systems.

An attack vector, or threat vector, is a way for attackers to enter a network or system. Common attack vectors include social engineering attacks, credential theft, vulnerability exploits, and insufficient protection against insider threats. A major part of information security is closing off attack vectors whenever possible. Essentially, attack vectors are a list of attacks that can be executed by exploiting the attack surface.

There exist differences between practice and theory. Some hypervisors may have relatively low surface attack vectors in theory, but there are a few emerging real-world attacks targeting hypervisors – e.g., the use of rootkits and covert channel calls. A hypervisor (its isolation mechanism is overpassed) compromised by a side-channel attack is likely to leak information outsourced to the cloud. Similarly, *layer spoofing*[27] has also been encountered, mapping to a new threat in the Blue Pill rootkit[28].

Some potential attack vectors in CC are as follows (Turnbull and Shropshire, 2013[29]) while a comparison of attack surface of each cloud service/delivery model is shown in Table 3:

- Redirecting Data Flows Using Firewall Ports

The first attack vector aims at data that is about to be transmitted to other computers. This attack vector refers to a method or pathway that an attacker can use to gain unauthorized access to a system or network. Redirecting data flows by manipulating firewall ports could be part of a broader strategy to intercept, exfiltrate, or manipulate data, making it a potential vector for an attack.

After escalating user privileges, the firewall could be amended so that data are redirected to the rootkit before they leave the host. This is accomplished by changing port rules in the firewall. The port of choice is usually number 22, one amongst a handful of open ports defined by the default configuration of ESXi. Other ports would include but are not limited to; port 443 (HTTPS access, ws-management), port 902 (host to host access, heartbeats), and port 5989 (CIM with HTTPS). Port 22 is used for the incoming Transmission Control Protocol (TCP) of the secure shell (SSH) server.

---

[27] Layer spoofing is a cyberattack technique where attackers forge or manipulate network packets at different layers of the OSI (Open Systems Interconnection) model to deceive cloud security mechanisms and gain unauthorized access, disrupt services, or bypass authentication.

[28] Blue Pill is the codename for a rootkit based on x86 virtualization. Blue Pill originally required AMD-V (Pacifica) virtualization support, but was later ported to support Intel VT-x (Vanderpool) as well. It was designed by Joanna Rutkowska and originally demonstrated at the Black Hat Briefings on August 3, 2006, with a reference implementation for the Microsoft Windows Vista kernel. The Blue Pill concept is to trap a running instance of the operating system by starting a thin hypervisor and virtualizing the rest of the machine under it. The previous operating system would still maintain its existing references to all devices and files, but nearly anything, including hardware interrupts, requests for data and even the system time could be intercepted (and a fake response sent) by the hypervisor. The original concept of Blue Pill was published by another researcher at IEEE Oakland in May 2006, under the name VMBR (virtual-machine based rootkit)

[29] Turnbull, L., Shropshire, J., 2013. "*Breakpoints: an analysis of potential hypervisor attack vectors*". In: Proceedings of the Southeastcon, IEEE. 2013. IEEE.

SSH is an excellent target because it is an internal agent within the hypervisor that communicates with the external network by passing packets through the firewall. Thus, it is here where an attack becomes viable.

- Infiltrating APIs

The next vector of attack highlights the exploitation of the application programmable interface (API) for malicious intent towards metadata. An API is a communication gateway connecting one application to another and allowing the translation of the data interchanged. An agent will communicate to another agent through an API and pass data back and forth, regardless of the programming language barrier. It is possible to intercept metadata as, for example, it is passed between applications in ESXi[30] by taking advantage of the application structure. The ESXi hypervisor, developed by VMware, is a widely used platform for virtualization in enterprise environments. However, like any complex software, it can have vulnerabilities that attackers may exploit, especially if hypervisor access is not restricted.

API infiltration attacks are a critical threat to cloud environments, as APIs serve as the backbone of cloud-based services, enabling communication between applications, users, and cloud platforms. When APIs are exploited, attackers can compromise data, manipulate cloud services, and even escalate privileges to gain deeper access to the environment.

Attackers identify public or private APIs exposed to the cloud environment. This involves enumerating endpoints, discovering weak authentication mechanisms, and looking for unprotected resources, due to vulnerabilities, like: Insecure Authentication and Authorization, Misconfigured Permissions, Lack of Input Validation, Exposed Endpoints, Poor Encryption, Lack of Logging and Monitoring, etc.

- Hooking Library Calls

Hooking library calls can be used as an attack vector. This technique involves intercepting and modifying the behavior of function calls in a program's libraries, which can be exploited for malicious purposes. Transitioning to a deeper level brings about a third vector of attack, hooking library calls. The concept of hooking is the ability to intercept, analyze, and customize a system or library call, before expediting it back to the call's originally intended target, being that of either the system or library file containing functions. The functions are scripts of code that produce desired actions, such as printing to another script or going as far as interacting with the kernel. Calls cannot be intercepted without first preloading a custom library prior to runtime. However, this can be achieved through configuring a system file to store a custom library file, and then having the library file preloaded to the system.

CC is based on the Internet and normally exposes its resources over the Internet. These resources can be categorized into three types according to the cloud layer that they concern (Bouayad et al., 2012[31]) (Fig. 10):

[30] VMware ESXi (formerly ESX) is an enterprise-class, type-1 hypervisor developed by VMware, a subsidiary of Broadcom, for deploying and serving virtual computers.

[31] Bouayad, A., et al., 2012. “*Cloud computing: security challenges*”. In: Proceedings of the Information Science and Technology (CIST), 2012 Colloquium in. 2012. IEEE.

- In SaaS model – Maps mainly to interface types that can be potentially supported: web UIs, GUIs, CLIs and RESTful APIs.
- In PaaS model – Web services: SOAP, REST, Runtime Engine and RPC protocols
- In IaaS model – VMs and storage services: Back storage, Object storage, File storage, Archive storage.

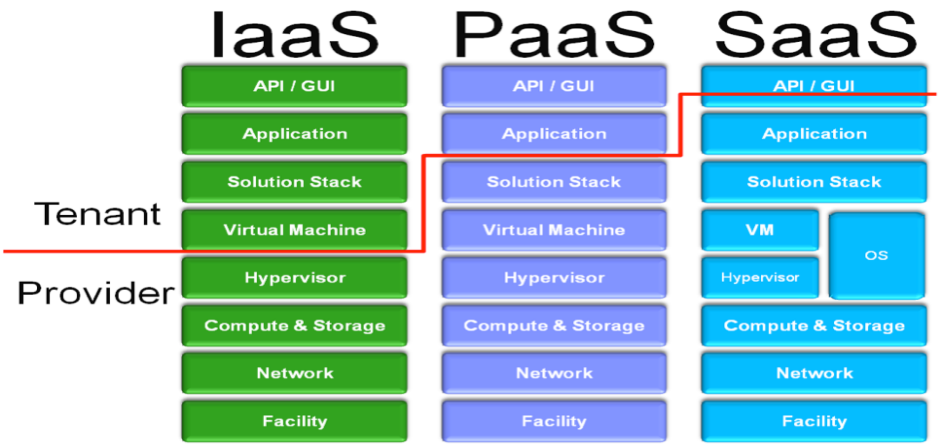

Figure 10: Resources of every cloud layer

Attacks surfaces in the cloud service models.

Attack surface	Attack Vectors		
	SaaS	PaaS	IaaS
Application level	Input/output validation	Runtime engine that runs customer's applications	virtual workgroups
Data Segregation	Unauthorized Access of Data	Data Service Portal	Multi-Tenancy and Isolation
Data Availability	Hosted Virtual Server	Network traffic	Virtual Network
Secure Data Access	Encryption/ Decryption Keys	Third Party Components	Cloud Multi-tenant Architecture
Data Center Security	Server based Data Breaches	Datacenter Vulnerabilities	Virtual Domain Environments
Authentication/Authorization	ID and Password	Client API Password Reset Attack	Poor Quality Credentials

Table 3. Attacks surfaces in the cloud service models (S. Iqbal et al./JournalofNetworkandComputerApplications74(2016)98–120, p. 103)

4.1.3 Attack vectors for cloud security on the SaaS cloud layer

Web applications represent a service at the upmost layer of CC. Such applications are dynamic services that pull data from different sources in a distributed cloud environment. Due to this feature, sometimes hackers insert scripts in the web pages of these applications by using form data. When these scripts get executed on the browser of the application’s customer, they cause unwanted behavior. SaaS application by using the specific means which is called the *script*. When these scripts get executed on the browser of the application’s customer, they cause unwanted behavior (e.g., the sending of critical cookie content of the attacker which could include user credentials or session keys). SaaS attack vectors targeting unauthorized access typically exploit vulnerabilities in authentication, configuration, API security, or multi-tenancy features. This more or less articulates an XSS attack.

Main attack vectors in SaaS cloud layer are:

- *Poor Input/Output validation:* Poor Input/output validation refers to ensuring that all data inputs to and outputs from an application are verified, sanitized, filtered and handled properly. In a SaaS cloud environment, poor input/output validation can introduce significant attack vectors, threatening the security and integrity of the service. SaaS applications are often multi-tenant, meaning they serve multiple users or organizations from a shared codebase and infrastructure. This makes input/output validation critical to prevent exploitation through injection attacks, buffer overflows, and data breaches.
- *Unauthorized access* of data: Occurs when attackers or unauthorized users gain access to sensitive information due to weak authentication, misconfigurations, or stolen credentials. This can lead to data breaches, identity theft, and compliance violations. Proper access controls, encryption, and multi-factor authentication (MFA) help mitigate this vulnerability in cloud environments. In SaaS environments, unauthorized access to data is a critical concern because SaaS platforms often host sensitive information, such as customer records, financial data[32], and intellectual property.
- *Attack vectors for Hosted virtual services*: Include vulnerabilities in hypervisors, insecure APIs, and weak VM isolation. Threats like hyperjacking, side-channel attacks, and misconfigurations can lead to unauthorized access or data breaches. Strong access controls, hypervisor security, and network segmentation are essential to mitigate these risks in cloud environments. Hosted virtual services[33] in SaaS environments provide databases, and file storage. While they bring efficiency and scalability, they also introduce specific attack vectors that can compromise data integrity, availability, and confidentiality. Common attack vectors for hosted virtual services in SaaS environments, are ‘Session Hijacking’ and ‘Account Hijacking’.
- *Attack vectors for Encryption/Decryption Keys:* Include weak key management, improper storage, insider threats, and brute-force attacks. If keys are exposed, encrypted data becomes vulnerable. Threat actors may exploit misconfigurations or steal keys from memory. Using hardware security modules (HSMs), key rotation, and strong access controls help mitigate these risks. Encryption and decryption keys are critical to securing data in SaaS environments. These keys protect sensitive data in transit and at rest, ensuring confidentiality and compliance with privacy regulations. However, if these keys are poorly managed or compromised, attackers can bypass encryption mechanisms, leading to significant data breaches. Common

[32] While financial data breaches often involve privacy violations, some cyberattacks on financial institutions do not necessarily steal personal information but rather manipulate transactions or disrupt banking operations.

[33] *Hosted virtual services* refer to cloud-based or remote computing services that allow businesses and individuals to access and manage resources over the internet. These services eliminate the need for on-premises infrastructure by providing scalable and flexible computing environments.

attack vectors associated with encryption and decryption keys in SaaS environments are: 'Key Exposure in API'[34] and 'Key Exposure in Logs'[35].

- *Server-based data breaches:* Occur due to misconfigurations, weak authentication, or unpatched vulnerabilities, allowing attackers to access sensitive data. Threats include SQL injection, malware, and unauthorized access. Implementing strong encryption, access controls, and regular security audits help mitigate these risks. In SaaS environments, server-based data breaches can have a severe impact on the security and privacy of users' data. Since SaaS providers typically handle the storage and processing of sensitive data on their infrastructure, a breach in their servers could expose that data to unauthorized parties. Here is some key attack vectors associated with server-based data breaches in SaaS environments: 'Insecure APIs' and 'Data Exfiltration Through Server Logs'[36].
- *ID and Password:* These vulnerabilities arise from weak credentials, reuse, phishing, and brute-force attacks. Compromised credentials allow unauthorized access to sensitive systems. Implementing strong passwords, multi-factor authentication (MFA), and password management policies help mitigate these risks in cloud environments. ID and password attacks are a critical security concern because they provide the primary means for users to authenticate and access applications. Weaknesses in how IDs and passwords are managed or protected can allow attackers to compromise user accounts and gain unauthorized access to sensitive data or services. 'Brute Force Attacks', 'Credential Stuffing', 'Phishing Attacks' and 'Password Reset Exploits', are some of the common vector attacks.
- *Hooking library attacks* (like 'API hooking'[37], 'IAT hooking'[38], 'Inline hooking'[39], etc) involve injecting malicious code or libraries into an application or system to intercept, manipulate, or monitor function calls, API requests, or data exchanges. Exploit vulnerabilities by intercepting function calls, keystrokes, or API communications in applications. Attackers can manipulate data, steal credentials, or alter system behavior. Preventing these attacks requires code integrity checks, secure coding practices, and endpoint security solutions. In the context of cloud environments, these attacks can be particularly dangerous as they allow attackers to exploit privileged processes, steal sensitive data, or manipulate the behavior of cloud-hosted applications. The attacker penetrates the application's server and it's done through other means like malicious input data leading to buffer/heap overflows

[34] *Key exposure in APIs* refers to the unauthorized disclosure or compromise of API keys, tokens, or credentials that allow access to APIs.
[35] *Key exposure in logs* refers to the unintended logging or storing of sensitive data, such as API keys, tokens, passwords, or other credentials, in system or application logs
[36] *Data exfiltration through server logs* is a covert method of extracting sensitive information from a server by exploiting logs
[37] Intercepts system or application API function calls to alter their behavior.
[38] Modifies the IAT (Import Address Table) table to replace a function pointer with a malicious one.
[39] Directly modifies binary instructions in memory to alter execution flow

4.1.4 Attack vectors for cloud security on the PaaS cloud layer

The PaaS cloud layer, or Platform as a Service, has a big role in being responsible for delivering the software execution environment that the PaaS customers require in order to run their applications seamlessly without needing to purchase any physical server, storage system, and networking equipment. In this regard, it is important to note that *security is considered of equal importance* in terms of the most essential feature, which is platform functionality (i.e., that a fully functional development and runtime environment is offered to the client to develop and run applications, and such an environment can be accessed via particular well-defined and implemented ways).

Correspondingly, it is a fundamental responsibility of PaaS providers to make sure that they do enforce strong security mechanisms (such as encryption mechanisms). Encryption is important when data or the actual code is transferred to the PaaS environment. Further, the PaaS providers become responsible for securing runtime engines from attackers. The programming/development stack (which supports developers and supply programming tools) that the PaaS vendors provide should also be secure from malicious threats. Multi-Tenancy (Rodero-Merino et al., 2012[40]) is another major attack vector in the PaaS cloud layer that covers two layers of virtualization: hardware and OS virtualization. In the latter case, OS virtualization enables us to offer different virtual platforms (in the form of containers) to different customers. The PaaS model has been developed in such a way that it allows multiple users to access PaaS services simultaneously; however, this feature also opens the possibility that a misbehaving user could leverage this access in many ways, affecting the normal functioning of the PaaS container / virtual platform. APIs are also supplied by the SaaS service to the customers via which they can access the offered platform. So, API security is a rather critical element in the PaaS landscape.

Main attack vectors in PaaS cloud layer are:

- *Vulnerable runtime engine that runs customer's applications:* In PaaS environments, the runtime engine is the core component that executes customer applications. It abstracts the underlying infrastructure and provides a managed environment in which developers can deploy, run, and scale applications without worrying about hardware or lower-level configuration details. However, since the runtime engine interacts with customer code, data, and external systems, it can suffer from several attack vectors. Attackers who exploit vulnerabilities in the runtime engine or misconfigurations (of the runtime engine) in the PaaS environment can compromise the security of the entire application stack, leading to data breaches, system outages, or unauthorized access to sensitive information. A common attack vector in this case is 'Code Injection'[41] (code injected in the runtime engine).
- *Data Service portal:* The data service portal is the interface or platform used by customers to interact with their data services. These portals typically provide functionality for managing databases, storage, analytics, and other data services.

[40] Rodero-Merino,L.,et al.,2012., "*Building safe PaaS clouds: a survey on security in multitenant software platforms*", Comput. Secur. 31(1), 96–108.

[41] *Code injection* is a type of attack where an attacker inserts malicious code into a program or system, which is then executed by the application

Since data is often a critical asset for organizations, the data service portal becomes a prime target for attackers looking to steal, modify, or manipulate sensitive data. A common attack vector in this case is ‘Cross-Site Scripting (XSS)’.

- *Network traffic:* Network traffic plays a central role in communication between the client applications, the PaaS infrastructure, and various other services or APIs. Since network traffic (between applications) can carry sensitive data, misconfigurations, vulnerabilities, and inadequate security measures can lead to a wide range of attack vectors. If an attacker can intercept or manipulate network traffic, they can potentially compromise the confidentiality, integrity, and availability of the platform and applications running on it. ‘MiTM’, ‘DDoS’, Session Hijacking’ and ‘Eavesdropping on Unencrypted Network Traffic’ are some of the common attack vectors
- *Third party components:* In a PaaS environment, third-party components (such as libraries, APIs, frameworks, and external services) are commonly integrated to provide additional functionality or ease of use. While these components can add significant value, they also introduce additional attack vectors, as vulnerabilities or misconfigurations in third-party software can potentially compromise the entire PaaS platform. A common attack vector in this case is ‘Data Leaks from Third-Party Integrations’[42].
- *Attack vectors targeting platform API password:* API-based password reset functionality is commonly used to allow users to reset their passwords securely. However, this feature, if not properly implemented and secured, can present significant attack vectors. Malicious actors may exploit vulnerabilities in the password reset process to gain unauthorized access to user accounts or compromise the platform. Some vector attacks in this case are: ‘Brute-Forcing Password Reset Tokens’, ‘Email or Phone Number Enumeration’ and ‘Reset Link Sent to an Attacker-Controlled Email’.

4.1.5 Attack vectors for cloud security on the IaaS cloud layer

The virtual machine monitor or hypervisor works either on top of OS or takes the role of a mini-OS. This technology is common to the IaaS cloud to support hardware virtualization technologies. Further, this technology is accompanied by APIs so that administrative operations may be performed (creating, deleting, staring, stopping objects, and writing Query programs to create or modify databases, forests[43], App Servers, and perform all kinds of administrative and configuration tasks).

[42] *Data Leaks from Third-Party Integrations* occur when external services or APIs connected to a cloud environment expose sensitive data due to weak security measures, misconfigurations, or vulnerabilities

[43] The term "*cloud forests*" in the context of Infrastructure as a Service (IaaS) doesn't have a widely recognized technical definition. However, it can be understood metaphorically or as an organizational analogy for complex, interconnected environments in the IaaS layer. Their characteristics are: Complexity and Layered Architecture, Resource Interdependence and Dynamic and Evolving Nature. As real-world Examples of Cloud Forests in IaaS, we can say that in multi-cloud environments, organizations often use multiple cloud providers (e.g., AWS, Azure, GCP) to diversify resources and improve resilience. This creates a forest of interconnected cloud ecosystems.

The addition of the hypervisor causes an increased attack surface. The reason is that there are many auxiliary attack methods that could be exploited over APIs or within different channels of communication like sockets, via different kinds of items, such as input strings. All of these give their own possibilities of exploitation (Szefer et al., 2011[44]).

Main attack vectors in IaaS cloud layer are:

- *Virtual workgroups:* In an IaaS environment, virtual workgroups refer to groups of virtual machines (VMs) or containers that work together to perform tasks or host applications. These workgroups are often used for tasks like load balancing, high availability, and scaling across a distributed architecture. However, virtual workgroups introduce multiple potential attack vectors that malicious actors can exploit if proper security measures are not in place. Some attacks are: ‘DDoS Attacks Targeting Virtual Workgroup Resources’ and ‘Cloud IAM Roles’ & Privilege Escalation’[45].
- *Datacenter (Infrastructure) vulnerabilities:* The underlying infrastructure often resides within a datacenter controlled by the service provider or a third-party cloud provider. The security of this datacenter plays a critical role in the overall security of the IaaS platform. If attackers exploit vulnerabilities in the datacenter’s infrastructure, it could lead to data breaches, denial of service, and compromise the entire IaaS environment. Datacenter vulnerabilities can stem from physical, network, or operational weaknesses. The attack vectors associated with these vulnerabilities affect not only the IaaS platform itself but also the applications and data hosted within it. ‘Physical Security Attacks’ and ‘DDoS Attacks on Datacenter Infrastructure’, are some such vector attacks.
- *Multi-Tenancy and tenant/resource isolation:* Multi-tenancy refers to the use of a shared infrastructure that supports multiple tenants (i.e., different organizations or users) within the same physical or virtualized environment. While multi-tenancy provides cost efficiency and resource sharing, it also introduces specific security concerns, particularly around isolation between tenants. If isolation between tenants is not properly enforced, attackers may exploit these weaknesses to access resources or data belonging to other tenants. ‘Cross-Tenant Data Access’ and ‘Side-Channel Attacks’ are some of the common vector attacks in this category.
- *Virtual Network:* Virtual networks play a crucial role in managing communication between different virtual machines (VMs), storage systems, and other network components. These virtual networks provide isolated, flexible, and scalable

[44] Szefer, J., et al., “*Eliminating the hypervisor attack surface for a more secure cloud*”, In: Proceedings of the Proceedings of the 18th ACM conference on Computer and communications security. 2011. ACM

[45] Overly permissive IAM (Identity and Access Management) roles allow attackers to escalate privileges and gain administrative control over virtual workgroups. For example, Amazon Elastic Compute Cloud (Amazon EC2) is a web service that provides resizable computing capacity—literally, servers in Amazon's data centers—that you use to build and host your software systems, and an AWS IAM policy which grants full access to all EC2 instances, might be:

```
{
  "Effect": "Allow",
  "Action": "ec2:*",
  "Resource": "*"
}
```

. A compromised user account enables the attacker to delete instances, steal data, or deploy malware.

networking environments that allow tenants to configure, manage, and control their networking resources. However, the complexity and shared nature of the infrastructure in IaaS environments also create several potential attack vectors that malicious actors can exploit to gain unauthorized access to resources, disrupt services, or eavesdrop on network traffic. Some common vector attacks are: ‘VM Escape’, ‘Side-channel attacks’, ‘MiTM attacks’ and ‘ARP Spoofing’ (Attackers send falsified Address Resolution Protocol messages to redirect traffic within a virtual network).

- *Virtual Domain Environments:* Virtual domain environments typically are used to control IT devices and users through a directory service that are provisioned and managed within a virtualized infrastructure. These virtual domains are created to provide tenants with a secure, dedicated environment within the shared physical infrastructure. However, various attack vectors exist that can compromise the integrity of virtual domains if not properly secured. ‘Hypervisor Attacks’, ‘Data Isolation Failures’ and ‘Lateral Movement Between Virtual Domains’[46] are some common attacks.
- *Poor quality credentials:* In an IaaS (Infrastructure as a Service) environment, the use of poor-quality credentials is a critical security risk. Credentials, such as usernames, passwords, API keys, and access tokens, are essential for authenticating and authorizing users, systems, and services to access cloud resources. When credentials are weak, poorly managed, or compromised, attackers can easily gain unauthorized access to sensitive infrastructure, leading to potential data breaches, service disruptions, and other security incidents. Some of the common attacks are: ‘Brute Force Attacks’, ‘Credential Stuffing’, ‘Password Spraying’ and ‘Phishing’.

4.2 Attacks, vulnerabilities and threats on Cloud Computing privacy

4.2.1 Cloud privacy vulnerability classification

Cloud vulnerabilities are weaknesses or gaps in a cloud computing environment that attackers can exploit to gain unauthorized access, steal data, or disrupt services. According to Forrester, the top 35 data breaches in 2022 accounted for 1.2 billion compromised customer records. Two different kinds of attacks are launched on authentication mechanisms in clouds. The access control interfaces are vulnerable to *signature wrapping* and advanced *cross site scripting (XSS) techniques* (Hussain et. al., 2017)[47]. In a signature-wrapping attack, an attacker modifies a signed message in such a way that the signature remains valid, while altering the meaning of the message. This can lead to unexpected repercussions, such as unauthorized access to sensitive information or funds. A *cross-*

[46] *Lateral movement between virtual domains* refers to the ability of an attacker or user to move or pivot across different virtual networks or domains within a larger networked environment. In the context of cybersecurity, it's often used to describe the technique employed by cyber attackers once they have breached one domain or system. They then attempt to gain access to additional virtual domains (or network segments) by exploiting vulnerabilities, misconfigurations, or weak security measures in the environment.

[47] Hussain S., Fatima M., Saeed A., Raza I., Shahzad R., “*Multilevel classification of security concerns in cloud computing*”, Applied Computing and Informatics, Volume 13, Issue 1, 2017, Pages 57-65, ISSN 2210-8327, https://doi.org/10.1016/j.aci.2016.03.001

site scripting (*XSS*) attack is a type of injection *attack* where malicious code is injected into a web page or application.

As companies increase their use of cloud hosting for storage and computing, the risk of an attack on their cloud services increases because the cloud presents a highly dynamic and distributed landscape with an enormous attack surface. As noted in the CrowdStrike 2024 Global Threat Report, there was a 75% increase in cloud environment intrusions in 2023. The report also revealed a 110% spike in cases involving cloud-conscious threat actors, which are threat actors that are aware of the ability to compromise cloud workloads and use this knowledge to abuse features unique to the cloud. Oftentimes, threat actors obtain valid credentials to access a victim’s cloud environment to then further extend their attack, typically leveraging tools that the organization has approved.

Besides, cybersecurity breaches go beyond cloud security and privacy attacks, affecting various industries and digital infrastructures. Many cyber threats target systems, networks, and devices without directly breaching cloud environments or stealing personal data. These attacks can still cause severe financial and operational damage. Poor cloud vulnerability management can cause reputational damage if customer data is compromised, leading to loss of business. Here is a categorization of cloud vulnerabilities affecting privacy.

4.2.1.1 Cloud misconfigurations

Cloud misconfiguration is one of the most common vulnerabilities organizations face. Misconfigurations can range from excessive account permissions to insecure backups. They are often caused by the speed of deployment (meaning mistakes can be made if the deployment process is designed and implemented in a rapid manner), limited knowledge of good practices (including security ones), or a lack of comprehensive visibility into the cloud infrastructure. Cloud misconfigurations, such as exposed storage buckets or improperly set access controls, directly compromise cloud privacy by allowing unauthorized access to sensitive data. These vulnerabilities often result from human error or lack of security awareness, leading to data breaches, regulatory violations, and erosion of user trust in cloud services

4.2.1.2 Insecure APIs[48]

APIs are widespread in modern software development, as they are used in microservices, applications, and website backends. They must handle legitimate requests received from mobile devices, applications, webpages, and third parties — as well as illegal requests from bots, spammers, and hackers. This is why having a secure API is critical to reducing your attack surface. APIs and UIs exhibit various vulnerabilities, such as weak authentication, lack of encryption, improper session management, insufficient input validation, outdated software, and overly permissive access controls. The effects of insecure interfaces in cloud systems can be significant, depending on the nature of the system and the security measures in place. These are usually technical, operational, financial and reputational. *Technically*, they expose vulnerabilities that can be exploited, leading to

[48] Insecure APIs, have also been referred as an example of ‘*Hosted virtual services*’ attack vector in SaaS layer (Section 4.1.3) and in Section 4.1.1.1 as APIs exploitation (SaaS vulnerabilities) and in Section 4.1.1.2 its usage, as a PaaS vulnerability.

data breaches, unauthorized access, and service disruptions. *Operationally*, such weaknesses can disrupt workflows, increase downtime, and strain IT resources as teams scramble to patch issues. *Financially*, the costs of breaches, regulatory fines, and lost business can be substantial, impacting profitability. *Reputationally*, insecure APIs erode customer trust, damage brand credibility, and can result in long-term loss of clients. Collectively, these effects highlight the critical need for robust API security measures to safeguard cloud systems and maintain operational and financial stability.

Insecure APIs in cloud systems can expose sensitive data to unauthorized access, enabling attackers to exploit vulnerabilities and breach privacy. Poorly designed or unprotected APIs may leak credentials, personal information, or confidential business data, leading to compliance violations, reputational damage, and loss of user trust in cloud-based services.

4.2.1.3 Lack of visibility

As the use of cloud services increases, so does the scale of your infrastructure. When companies are using thousands of instances of cloud services, it can be difficult to track how they are all connected or see which ones are running in production at any given time. Visibility into the state of the entire infrastructure must be easy and convenient to access. Lack of cloud infrastructure visibility is a major issue that makes it difficult to take action on a threat, since finding the source of a vulnerability is very difficult. Both security and DevOps teams must maintain continuous visibility into their real-time cloud security posture, since security teams have insight into risk and DevOps teams have ownership of cloud resources and applications. Lack of visibility in cloud environments makes it difficult to monitor data access and usage, increasing the risk of unauthorized exposure or breaches. Without proper oversight, sensitive information can be accessed, shared, or misused unnoticed, compromising privacy and leading to regulatory non-compliance, reputational harm, and loss of customer trust.

4.2.1.3.1 Shadow IT

One of the main reasons that organizations lack visibility in their cloud infrastructure is the use of shadow IT, which refers to the practice of creating cloud resources or any other digital asset without proper approval from the IT department. Shadow IT is prevalent when companies experience rapid growth, as employees may bypass the approval process to minimize disruptions to their day-to-day operations. Shadow IT presents security risks because unauthorized assets are often not properly secured due to negligence — for example, employees may keep default passwords or misconfigurations because these assets were created outside of an approved, security-aware process. Shadow IT, the use of unauthorized cloud services or applications, bypasses organizational security controls, creating blind spots. This can lead to improper handling of sensitive data, increased exposure to breaches, and compliance violations. Without oversight, privacy risks escalate, as data may be stored or shared insecurely, undermining trust and regulatory adherence.

4.2.1.4 Poor access management

Having insecure and poor access management is a common risk in cloud systems. At a high level, it occurs when users or services have access to resources that they should not be able to access

and/or do not need. Poor access management can lead to adversary exploitations like account hijacking. Poor access management in cloud systems allows unauthorized users to access sensitive data, increasing the risk of breaches and privacy violations. Overly permissive permissions, shared credentials, or lack of role-based controls can lead to data exposure, regulatory non-compliance, and reputational damage, undermining trust in the organization's ability to protect user information.

4.2.1.5 Zero-day vulnerabilities

A *zero-day* is a *vulnerability* in software or hardware that is typically unknown to the vendor and for which no patch or other fix is available. These types of vulnerabilities are therefore undetectable by many antivirus software solutions or other signature-based threat detection technologies. Once they have exploited the vulnerability, attackers might attempt to exfiltrate sensitive data, perform remote code execution, or block legitimate users from accessing their cloud services. Zero-day vulnerabilities in cloud systems are unknown flaws exploited by attackers before patches are available. These can lead to unauthorized access, data breaches, and exposure of sensitive information, compromising user privacy and resulting in regulatory penalties, reputational damage, and loss of trust.

4.2.1.6 Human vulnerabilities

According to the Thales Global Cloud Security Study, human actions were responsible for 44% of cloud data breaches reported incidents. These errors can take many forms, including misconfigurations and access management issues. Many of these errors are caused by limited knowledge about security best practices or poor strategic planning. Human errors are a consequence of human vulnerability, so they do not actually constitute vulnerability. Human vulnerability is, for example, a poorly skilled or educated employee who is not aware of the best practices and can be exploited by social engineering attacks. Human vulnerability, such as phishing or poor password practices, often compromises cloud privacy. Employees may inadvertently expose credentials or sensitive data, enabling unauthorized access. This human error can lead to data breaches, regulatory violations, and reputational harm, undermining trust in cloud security.

4.2.1.7 Vulnerabilities in data leakage prevention systems

To implement a successful DLP solution, specific challenges, as discussed in the following subsections, must be considered and addressed adequately.

(I) *Leaking channels*

For data “in use” and “at rest”, confidential data can be leaked through channels, such as USB ports, CD drives, web services and printed documents. Data leaks through USB ports and CD drives can be mitigated through host DLPSs, but this is not enough since other leak channels, such as emails, are always available.

(II) *The human factor*

It is always difficult to predict human behavior because it is influenced by many psychological and social factors. Many human actions are affected by subjectivity in making decisions, such as defining the secrecy level of data, assigning access rights to specific users and calibrating a detection

threshold of a DLPS. In addition, conformity with the organization’s security policy is not always guaranteed, even with strict regulations and guidelines.

(III) *Improper definition and violation of access rights*

It is important for DLPSs to be able to distinguish between different users based on their privileges and permissions. Without a proper definition of access rights, DLPSs cannot decide whether the data is being accessed by a legitimate user. Therefore, access control systems play an important role in preventing data leakage.

(IV) *Encryption*

For network-based DLPSs, encryption is considered a major challenge. Network-based DLPSs attempt to identify copies of confidential data using various analysis techniques by also comparing them to the original data. Encryption is a way for data—messages or files—to be made unreadable, ensuring that only an authorized person can access that data. Encryption uses complex algorithms to scramble data and decrypt the same data using a key provided by the message sender. Encryption ensures that information stays private and confidential, whether it's being stored or in transit. Any unauthorized access to the data will only see a chaotic array of bytes. However, the use of strong encryption algorithms[49] makes it somehow difficult to analyze the data content.

(V) *Data Modification*

Some DLPs use data patterns and signatures to compare inspected traffic and original confidential data, to detect data leakage. The detection takes place when these patterns and signatures are matched or when a high degree of similarity is noticed. However, confidential data is not always sent as is. In fact, confidential data can be exposed to many types of modifications.

(VI) Scalability *and integration*

Like many other network security mechanisms, DLPSs too can be affected by the amount of data being processed. Whether the DLPS is host based, network based, or storage based, DLPSs should be able to process data without delaying the workflow. Factors, such as the inspected data size, the computational capabilities and the analysis technique used, should be fully investigated to run a scalable DLP system. Scalability and Integration are, essentially, indirect threats, as for example DoS, in the sense that the DLP becomes a bottleneck and can act as a “Trojan Horse” in terms of DoS as it could significantly lower the performance of the system in which, it is run. It could be also a medium via which DoS attacks could take place by knowing their existence.

(VII) *Improper Data classification*

DLPSs depend considerably on appropriate data classification. If the data is not classified into different levels, DLPSs will not be able to distinguish between confidential and normal traffic. In military and government sector applications, the use of data classification is common.

[49] Some encryption algorithms are: *Triple Data Encryption Standard*, *Advanced Encryption Standard*, *RSA Security*, *Blowfish*, *Twofish*. Also, started in 2016 as the NIST's Post-Quantum Cryptography Standardization project, four algorithms have been winnowed down from 69 submissions by cryptography experts in dozens of countries: *CRYSTALS-Kyber* (FIPS 203), designed for general encryption purposes, such as creating websites, *CRYSTALS-Dilithium* (FIPS 204), designed to protect the digital signatures used when signing documents remotely, *SPHINCS+* (FIPS 205) is also designed for digital signatures and *FALCON* is also designed for digital signatures and is slated to receive its own draft FIPS in 2024. The analysis of the above algorithms is out of the scope of the thesis.

4.2.1.8 Container vulnerabilities

Container vulnerabilities arise from misconfigurations, insecure images, or lack of isolation, exposing risks like privilege escalation, data breaches, or denial of service. Common issues include unpatched software, exposed ports, and weak access controls. Proper practices like using trusted images, regular updates, and tools like Kubernetes security policies can mitigate these risks.

1. **Misconfigured Containers:** One of the most common security issues with containers is misconfiguration. Containers often operate with elevated privileges or are assigned too broad access to resources, creating security gaps. For example, a container might have unnecessary root privileges or it could be allowed to have insecure communication between containers. Misconfigurations, such as open network ports, unencrypted communication, or improperly configured authentication mechanisms, can expose sensitive information. Misconfigured containers can become an easy target for attackers looking for access to valuable data in the cloud.
2. **Insecure Container Images:** Container images serve as the foundation for containerized applications. However, many container images are downloaded from public repositories, and some may not be vetted for security vulnerabilities. Attackers can exploit insecure or outdated container images by embedding malicious code, backdoors, or vulnerabilities within them. These vulnerabilities can lead to data leakage or unauthorized access to critical resources once the container is deployed. Without proper scanning and validation, the use of untrusted or insecure container images becomes a serious risk to cloud privacy. Please note at this point that a container image can be constructed from other base container images, so the latter can be utilized to spread security issues & vulnerabilities to the former. Thus, users should be careful not only about the new content (e.g., vulnerable dependencies/libraries) of new images to be constructed but also the current content of the base images exploited for this construction.
3. **Insecure Container Orchestration:** Containers are often managed using orchestration tools, such as Kubernetes, which automates container deployment, scaling, and management across multiple resources. However, improper configuration of orchestration tools or the use of outdated versions can create security vulnerabilities. If orchestration tools are not secured properly, attackers can potentially gain control of the orchestration platform itself and manipulate the containers or compromise sensitive data. For example, unauthorized access to Kubernetes clusters through weak authentication or exposed APIs could lead to malicious activity, including the compromise of cloud privacy.
4. **Weak Container Network Security:** Containerized applications often interact with each other over a network. If the container network is not properly secured, attackers can exploit vulnerabilities in communication channels between containers. Lack of encryption, poorly configured network segmentation, or open network ports can expose sensitive data in transit. Attackers may intercept or tamper with unencrypted communications, leading to data breaches or the compromise of privacy.

4.2.2 Cloud data privacy vulnerabilities

4.2.2.1 Cloud Data Privacy vulnerabilities in cloud layers

Although countermeasures can be used to secure cloud data in terms of data confidentiality, data integrity, and data availability and support secure data sharing in groups and privacy protection, there exist vulnerabilities in cloud environments that could impact their effectiveness. The major of these vulnerabilities are as follows:

(I) Security data vulnerabilities and concerns related to SaaS

A. *Unprotected sensitive data or information flow over the network*: In a SaaS environment, most of the security concerns arise based on the sensitive information flow over the network. Most of the attackers target this flow and attempt to flood the requests or extract/intercept the information exchanged (information is also altered to cause undesired effects, through XSS, SQL injection, XML signature wrapping, etc)

B. *Issues based on Resource allocation*: In a SaaS environment the cloud consumers use the services/applications offered by a CSP. These consumers do not know about the location of resources, so they do not know the location of their data that is processed and stored by the respective SaaS services. By conforming to privacy laws across different countries, CSPs share personal consumer data with countries that might not have proper security provisions for them and in some cases the CSP or the consumer would have to pay fees for not conforming to law. Even more, an SLA might indicate that it is the fault of the consumer, although he/she does not control the location of his/her data.

C. *Problems-in-Multitenancy*: In a SaaS cloud environment, such problems are related to non-proper isolation of data (at the application level), so data of some users are exposed to other users. Due to applying multitenancy practices potentially across different cloud layers, the information or data of various users/customers resides in the same location This increases the probability that user data is leaked due to various kinds of attacks, like side-channel ones.

(II) Security data vulnerabilities and concerns related to PaaS

A. *Data Loss.* As a concern, we can consider that data stored in PaaS environments may be lost due to accidental deletion or malicious activity. Known vulnerabilities include:

- Lack of regular backups or inadequate backup strategies,
- Insufficient disaster recovery plans,
- Shared responsibility model misunderstandings (e.g., assuming the provider handles all data protection).

B. *Inadequate Security Provisions for Shared Objects*: A PaaS environment relies on an identity model; CSP provides the security to infrastructure and the offered platform, and the PaaS customers have to protect their accounts based on their identity. There are no proper provisions for the objects shared, configurations, etc. In this way, inadequate provisions for securing objects (e.g., resources, data, and APIs) and configurations can create significant vulnerabilities. Regarding the lack of proper provisions for shared Objects, we refer to the following vulnerabilities:

- *Inadequate Access Control*: Shared objects, such as storage buckets, databases, or APIs, often lack granular access control, making them vulnerable to unauthorized access.

- *Overexposed Resources*: Default settings may allow overly permissive access, such as making resources public or accessible by all authenticated users.

- *Improper Data Isolation*: Shared environments (e.g., multi-tenant PaaS platforms) might fail to fully isolate resources between different tenants or projects.

- *Weak Encryption*: Data stored in shared objects may not be properly encrypted at rest or in transit, increasing the risk of unauthorized exposure.

- Also, there is *Weak Monitoring and Logging of Shared Resources* and *Lack of Automated Provisioning and Governance*. "Weak Monitoring and Logging of Shared Resources" refers to insufficient tracking and analysis of activities in shared Platform-as-a-Service (PaaS) environments. This gap can lead to undetected breaches, misuse, or unauthorized access, compromising data security. Without robust logs, identifying and responding to threats becomes challenging, increasing vulnerability to attacks. "Lack of Automated Provisioning and Governance" occurs when PaaS environments lack automated tools to manage resource allocation and enforce security policies. Manual processes are error-prone and slow, leading to misconfigurations, over-permissioned accounts, or non-compliance with security standards, creating exploitable gaps in data protection.

C. *Encryption Gaps*. As a concern we can consider that Data encryption is critical for protecting sensitive information, but gaps in encryption practices can leave data exposed. Known vulnerabilities include:

- Data transmitted or stored without encryption,
- Use of weak encryption algorithms or outdated protocols,
- Poor key management practices (e.g., hardcoded keys, lack of key rotation).

(III) Security data vulnerabilities and concerns related to IaaS

A. *Problems in Private Cloud access*: In an IaaS environment, handling tenant identity management in terms of authentication and authorization as well as handling administration are the major problems that might not have proper solutions. Indeed, since the lack of centralized authentication and multi-tenant isolation plus the shared responsibility confusion, etc, exist as challenges, in the private cloud access, it is required to balance security, usability, and scalability.

B. *Shared Technology Vulnerabilities.* As a concern, we can say that IaaS environments often rely on shared infrastructure (e.g., hypervisors, storage, and networks). Vulnerabilities in these shared components can affect multiple tenants. Known vulnerabilities include:

- Hypervisor vulnerabilities (e.g., VM escape attacks),
- Improper shared resource isolation leading to successful side-channel attacks,
- Lack of proper isolation between tenant environments.

C. *Virtual Machine (VM) Sprawl,* As a concern, the Uncontrolled creation and deployment of VMs can lead to security gaps and increased attack surfaces. Known vulnerabilities include:

- Unpatched or outdated VMs,
- Orphaned VMs that are no longer managed but still running,

- Lack of inventory and tracking of VMs.

4.2.3 Cloud data privacy threats in cloud layers

We classified the following threats for cloud privacy, taking also under consideration our proposed model in Ch. 4, section 4.1. The security provided for the cloud data in terms of data confidentiality, data integrity, and data availability, secure data sharing in groups and privacy protection faces some threats, the majority of which are as follows:

(I) Security data threats and concerns related to SaaS

A. *Data Residency and Sovereignty.* Data stored in SaaS applications may be subject to the laws and regulations of the country where it is stored, which can create legal and compliance challenges. The related concerns include:

- Lack of control over data storage locations,
- Cross-border data transfers that violate local regulations,
- Inability to ensure data sovereignty.

B. *Compliance and Legal Risks.* Storing and processing data in SaaS applications may introduce compliance challenges, especially for regulated industries (e.g., healthcare, finance). The related concerns include:

- Non-compliance with data protection regulations (e.g., GDPR, HIPAA, CCPA),
- Lack of transparency in data storage locations (e.g., cross-border data transfers),
- Inadequate logging and reporting for audit purposes.

C. *Data Retention and Deletion:* SaaS providers may retain data longer than necessary, or data may not be properly deleted when requested. The related concerns include:

- Lack of clear data retention policies,
- Difficulty in ensuring complete data deletion,
- Potential for data to be recovered or accessed after deletion.

(II) Security data threats and concerns related to PaaS

A. *Data breach*: PaaS is usually based on sharing of resources in terms of network, hardware etc. Security concerns rise on data breach if the hackers can obtain unauthorized access to the PaaS tenant accounts so as to obtain privileges to access the respective offered platform/environment such that they can disclose/leak data or hamper the integrity or availability of the application(s) running in the platform.

B. *Insider Threats.* Both malicious and negligent insiders (e.g., employees, contractors, or third-party vendors) can compromise data security in PaaS environments. The related concerns include:

- Excessive privileges granted to users or applications,
- Lack of monitoring and auditing of user activities,

- Insufficient background checks for employees with access to sensitive data.

C. *Insecure Third-Party Dependencies.* PaaS platforms often integrate with third-party services or libraries, which can introduce security risks. The related concerns include:

- Use of outdated third-party components,
- Lack of vetting for third-party services,
- Supply chain attacks target dependencies where cybercriminals infiltrate software providers or third-party vendors. Attackers compromise legitimate software updates to spread malware to end users, as seen in incidents like the SolarWinds attack. This method allows hackers to bypass traditional security defenses by exploiting trusted sources.

(III) Security data threats and concerns related to IaaS

A. *Permanent Data Loss*: Data loss due to unexpected disasters at data centers or malicious attackers targeting the cloud virtual and eventually physical machines by exploiting misconfigurations or improper resource isolation may cause severe security concerns.

B. *Remote Network threats*: Remote network threats exist in IaaS cloud layers because cloud infrastructure relies on internet-accessible networks, making it vulnerable to attacks. Misconfigurations, weak encryption, and inadequate monitoring further increase risks, allowing attackers to disrupt services, steal data, or gain unauthorized access to sensitive resources. These threats can disrupt services, expose sensitive data, and grant unauthorized access, emphasizing the need for strong encryption, firewalls, and intrusion detection systems.

C. *Unpatched Software and Security Misconfigurations.* Attackers exploit outdated software, default credentials, and misconfigured settings. Causes may be the Unpatched operating systems, Default security settings left unchanged (e.g., admin passwords) and Lack of regular security assessments. As an example, if the server has misconfigured security groups or firewall rules, such as leaving SSH port 22 open to the public, attackers can exploit these weaknesses to gain unauthorized access, deploy malware, or exfiltrate sensitive data.

Below, we categorize the main cloud privacy attacks on every cloud layer. Some of these attacks can be considered as “*Multilayer attacks*”. In cloud computing the term ‘multilayer attack’ refers to cyberattacks that can target two or more layers of the cloud architecture distinctly (see also Table 4). A ‘cross-layer attack’ can attack simultaneously multiple layers in order to fulfil its purpose. We refer, here, to multilayer attacks.

Table 4. Cloud attacks and possible defense (Salih B., et al., 2024)

Attack Name	Attack process	References	Defenses
CrossVM attack – Attacks on Confidential Cloud data	Uses Side channels and attacker places the malicious Virtual Machine on the Clients physical server location.	Ristenpart	Yes – Nohype TCCP Co-residence
Malicious SysAdmin Attack on Confidential Cloud data	Authorized system admin on the cloud provider may access customer memory of users VM.	B.D. Payne	Yes PDP
Data loss or Manipulation Attack- Attack on Cloud Data Integrity	Due to large amounts of cloud storage in distrusted servers. Attacks can be initiated using data migrations, loss of data by taking adavantage of cloud owners' loss of control	G.Ateniese	Yes HAIL
Dishonest Computation – Attack on Cloud Data Integrity	Due to mis- configurations, outdated and vulnerable code at cloud servers.	A. Baliga	Yes Re-computation
Flooding attack – Attack on Cloud Data Availability	Due to huge number of requests makes the service unavailable causing to Denial of service attacks	M. Jensen	Trusted computing Yes Service Migration FRC-attack- Detection
SLA violation Attack on Cloud Accountability	Due to improper allocation of resources to cloud customer by cloud provider	A. Haeberlen	Yes Collaborative monitoring
Dishonest MapReduce Attack on Cloud Accountability	Due to mis- configurations in working machines cloud results in accurate results	W. Wei	Accountable VM Yes Accountable Mapreduce
Data Privacy and Computation privacy attack on Cloud Privacy	Due to personal information leakage using advanced hacker techniques	M. Mowbray	Yes Privacy Manager Privacy-as-a- service

4.2.4 Cloud privacy attacks on the SaaS cloud layer

Based on the research and findings carried out by Gartner, Software as a Service, commonly known as SaaS, can be defined as the "software that's owned, delivered and managed remotely by one or more providers". In addition to the above definition, Gartner (Council, 2006), offers an estimate, which shows more than 75% of companies' information technology budgets is spent on the purposes of buying, operating, and maintaining software systems and infrastructure.

In 2011 (Barron et al., 2013), researchers from the Ruhr-University Bochum found a security hole in the cryptography in Amazon's EC2 and S3 models. There is a flaw in the security protocol of web services that enables the attackers to bypass the security layer of digitally signed SOAP messages. As such, the control interfaces that are used to manage the cloud resources can actually be hijacked by the attackers. In this way, the attackers would be able to create, modify, and delete the machine images and alter the administrative password. The authors present an algorithm based on XML digital signature.

Since Software as a service (SaaS) allows users to connect to and use cloud-based apps over the Internet, users' data are stored in relational databases (include SQL Server, Oracle, MySQL, PostgreSQL, Spanner, and Cloud SQL) and in non-relational cloud databases which store and manage unstructured data, such as email and mobile message text, documents, surveys, rich media files, and sensor data. In this regard, one can securely outsource the respective data (in an encrypted form when in transit and stored - it is the responsibility of the SaaS provider to protect them), hence promising a high confidentiality and security level about the information concerned. However, basic questions emerge here, such as who owns the data (that are being produced when the customer interacts with the SaaS application), the process of data backup, accessibility to the data, the location of the data, general availability, and essential elements of identity management and authentication and access control to cover both authentication and authorization, among others.

The following cases of attacks are related to the SaaS layer and are analysed as follows:

4.2.4.1 Denial of service (DoS) and Distributed DoS (DDoS) attacks

A DoS attack is an intentional malicious attempt to render the system or network resources unavailable to users. Many millions of users share CC infrastructures, so that it is becoming more difficult to resolve this sort of attack as it might have a much greater impact when compared to single tenanted architectures (Khan et al., 2016[50]). Sophisticated attackers may attack the whole infrastructure, and they dedicate multiple resources to conduct this attack and bring down a targeted service or application. Denial of Service is typically accomplished by flooding the targeted machine or resource with superfluous requests in an attempt to overload systems and prevent some or all legitimate requests from being fulfilled. The range of attacks varies widely, spanning from inundating a server with millions of requests to slow its performance, overwhelming a server with a substantial amount of invalid data, to submitting a magnitude of requests with an illegitimate IP address.

(D)DoS attacks also pose a significant risk. These attacks overwhelm, also, *servers* or *networks* with massive traffic, causing service outages. At a different logic, the attack may focuses on the application layer or platform services, such as APIs, databases, or web applications hosted on the platform, making it as a multilayer attack at the PaaS layer, too. While they do not necessarily involve *data breaches*, they can lead to revenue loss and reputational damage for businesses.

A distributed denial-of-service (DDoS) attack is a malicious attempt to disrupt the normal traffic of a targeted server, service or network by overwhelming the target or its surrounding infrastructure with a flood of Internet traffic. DDoS attacks achieve effectiveness by utilizing multiple compromised computer systems as sources of attack traffic. Exploited machines can include computers and other networked resources such as IoT devices[51]. From a high level, a DDoS attack is like an unexpected traffic jam clogging up the highway, preventing regular traffic from arriving at its destination. Thus, a multi-level security strategy might be in effect in these cases.

As such, the Botnets play a major role in the dissemination (which takes the form of a DDoS and not just a DoS attack) of these attacks. They comprise compromised machines in on premise environments, machines in the cloud and IoT devices. More resource-intensive applications running at the cloud with low bandwidth may attract attacks if the attackers have such knowledge, so it will be easier to bring them down than other applications. As in CC, new weaknesses have been opened as the use of virtualized data centers and cloud services is increasing. Modern attackers do not need to attack the full infrastructure. An example is how Twitter was overwhelmed by DoS attacks in 2009.

SaaS stands for Software as a Service, and these offer their valued customers a range of vital services. If these services are made unavailable by the disruptive actions perpetrated by DDoS attackers, the SaaS customers will not get value for their money in respect to the service they have paid for, having an improper performance experience with that SaaS application. This rather

[50] Khan, S., et al., 2016. “*Cloud log forensics: foundations, state of the art, and future directions*”. ACM Comput. Surv. (CSUR) 49 (1), 7.

[51] Vulnerabilities in the Internet of Things (IoT) and control systems attacks expose smart devices to cyber threats. Hackers can take control of industrial machinery, home security systems, or even medical devices, causing severe consequences.

unfortunate situation explains why SaaS services tend to become very attractive targets for destructive DDoS attackers who wish to bring down services.

It is usually very difficult to detect and trace the identity of the attacker who has conducted a DDoS attack. This concern urgently needs to be overtaken drastically by a comprehensive study in the entire system, focusing on the use of robust intrusion detection and prevention mechanisms. This strategic approach is meant to handle the massively growing threat of DDoS.

A typical DDoS scenario is shown in Fig. 11 where the main entities involved in such an attack are depicted. The *handler* is probably the most high-volume server in which attack packets can be hidden easily. Agents are hosted in user/enterprise/cloud machines that are already infected. In this way, it is very difficult to track down the attacker. A victim could be a target in the SaaS application (e.g., a virtual or physical machine on which the application is running). Usually, attackers will try to place the handler software on a compromised router or network server that handles large volumes of traffic. This makes it harder to identify messages between the client and handler and between the handler and agents. In descriptions of DDoS tools, the terms “handler” and “agents” are sometimes replaced with “master” and “daemons”, respectively.

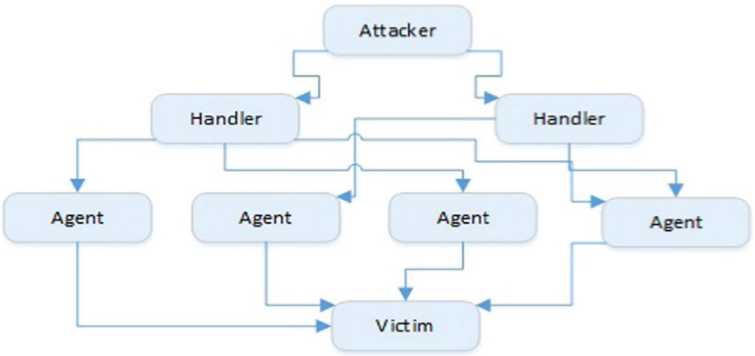

Fig. 11. A general diagram for distributed DoS

4.2.4.2 Authentication attack

For both users and CSPs, authentication is a significant security task. The primary intent behind authentication is to control the access of the data processing system on behalf of the authorized persons by conforming to the CSP’s policy. Essentially, a mechanism or the method of authenticating the system becomes an all-time target to the attackers in a strive to bypass authentication and have unauthorized access to that system.

As authentication attack techniques we can refer to the following: (a) Session hijacking is an attack on access control (*Session Hijacking attack* compromises the session token by stealing or predicting a valid session token to gain unauthorized access to a Web Server) and (b) Credential harvesting is another one (*Credential harvesting* is a cyberattack technique where cybercriminals gather user credentials — such as user IDs, email addresses, passwords).

For example,[52] the LastPass Company that stores and manages passwords in the CC environment indicated that their systems were affected from a successful hack (a side problem). However, no data was stolen, but the company encouraged their customers to take measures to

[52] For more real cases, regarding attacks on every cloud layer, the reader can refer to Ch. 4, Section 4.2.7

ensure that their data is secure. On the other hand, security experts found strange activities in their systems because more data was going out as compared to incoming data, but without data leakage.

It was a significant cybersecurity incident affecting one of the most popular password management services. LastPass stores encrypted user passwords and sensitive data, making the breach particularly concerning. Attackers gained access to LastPass's development environment by compromising a developer's account. This was achieved through a phishing attack or by exploiting a vulnerable third-party software component. Once inside, the attackers stole source code and proprietary technical information from LastPass's development environment.

Authentication verifies that a user is who they say they are. Authorization is the process through which an administrator grants rights to authenticated users, and the process of checking user account permissions to verify the user was granted access to those resources. It is important to authenticate for both users and CSP. The primary purpose of authentication is to allow only authorized persons to access the data processing system according to the cloud provider's policy. In essence, the mechanism and methods used to authenticate the system are frequent targets by attackers. For a certain class of cloud applications, many users are still using the simplest and unsophisticated mechanism of username and password to verify their account.

Typically, two kinds of authentication methods are offered by SaaS providers depending on their architecture: a *centralized authentication* and a *decentralized authentication* system.

4.2.4.3 XML signature wrapping attack

Attacks can be launched against SOAP messages targeting web services that operate in cloud environments (web services programmatically enable the consumption of the main functionality of a web application like a SaaS). We call these attacks "XML Signature Wrapping Attacks” (Qaisar and Khawaja, 2012).

We have already described that the SOAP messages become vulnerable to attacks. These are referred to as XML Signature Wrapping Attacks. The attacks on protocols by using XML Signature can also affect the CC services. The worst-case scenario occurs when these attacks even break the security architecture between the client/browser and the cloud/web service.

In this manner, an attacker can authenticate himself as a valid user while making several arbitrary web service requests. However, this case is somehow difficult to implement, because the attacker has to change the content of the message without being detected.

In addition to being utilized for communication between the client and server, this message contains structural information. Attackers use XML signatures to authenticate digitally signed SOAP messages that have been altered by them if the web server security protocol has any flaws (vulnerabilities such as SQL injection, cross-site scripting, and user session management flaws). In this respect, if the original signature is removed, and the previous content is retained and new content is inserted, the complete message is digitally signed, too. Of course, since the signature is produced via a different key, there will be an inconsistency with the original key, but if the client key is leaked, then the attacker can use a new “authorized” message. In this case, the unauthorized message alteration might not be detected by the web service, even if it applies to the UNWRAP method.

When the user submits a request via his machine’s browser, the server generates the SOAP message. Web services are called via SOAP by client programs. Having the SOAP request (as sent by the client) and the SOAP response as returned by the web service, in this method, the XML signature wrapping attack is prevented by employing the concept of XML digital signature on the SOAP message.

4.2.4.4 Cloud malware-injection attacks (SaaS case)

In malware injection attacks, which is generally an application-based attack category, attackers insert or inject malicious code into a web page of a (SaaS) application causing web pages (when users browsing them) to be infected and in turn these malicious web injections can be used to steal sensitive information, such as login credentials and personally identifiable information. The static files in web applications are usually controlled by the web server. But during injection, the dynamic content is usually affected that is usually drawn from underlying databases. Thus, there is a need to store the malicious script in the database such that it can then be loaded within a specific part of a dynamic page. When a user sends their request to perform an operation, attackers hack it and inject on it the malicious content. In this case, the code appears as a valid instance of services running on the cloud.

4.2.4.4.1 SQL injection attack

A major attack used to steal users' information (or data can be altered or erased, violating their integrity or availability) from a web application (like a SaaS) is SQL-injection. In this type of attack, the malicious code/script is unwittingly inserted into normal SQL code by the application when constructing this code, which changes its nature. When this code then executes, the malicious actors get access to data stored in the underlying database or can modify or delete it (Bhadauria et al., 2014[53]).

Most of the SQL injection attacks happen through user submitted string-based input: the first part contains the guess containing the information to securely terminate the original SQL command; the second part is the hostile code which the attacker wants to run on the underlying database.

4.2.4.4.2 Cross-site scripting (XSS attack)

Cross-site scripting or XSS is also amongst the most prevalent kinds of application layer hacking techniques. Since this injects malicious scripts into web contents, it is the kind of injection attack (Rodero-Merino et al., 2022[54]). As the cloud provides a shared environment, an attacker tries to inject malicious scripts in web programming languages like JavaScript, HTML, and VBSCRIPT into Dynamic Web Applications. Such scripts take the form of Browser Side Scripts as they become part of the application’s web page that is rendered by web browsers. Once this web page is loaded in a client machine through the browser, the malicious script is executed to realize its malicious intent, which can be the gathering of important/sensitive information from the client’s machine-like

[53] Bhadauria, R., et al., 2021. “*A survey on security issues in cloud computing*”. arXiv preprint arXiv:1109.5388, 2011

[54] Rodero-Merino, L. ,et al., 2022. “*Building safe PaaS clouds: a survey on security in multitenant software platforms*”. Comput. Secur. 31(1), 96–108.

session keys. Automated scanners can also be used to detect the most common vulnerabilities in addition to manual checking. In this case, multiple web application scanners are helpful to detect any vulnerability, such the Cross Site Scripting (XSS) and Structured Query Language (SQL) injection attacks.

XSS attacks employed for stealing or modifying user data through the insertion of malicious code; this code is injected into the web application in the form of user input. Hence, when users submit data using forms in the web page of a SaaS application, there are two main alternative cases for XSS realisation: (a) the users are malicious (but appear as normal application authenticated users) and include in their input data malicious scripts or (b) malicious actors intervene as an intermediary, intercepting the user form data and changing it to include the malicious script before it is handed over to the application.

Both Cross-Site Scripting and SQL Injection attacks are becoming serious threats for attacks. FireHost (FireHost, 2012) reportedly prevented over 64 million malicious cyber-attacks. Most of these attacks are cross-site scripting attacks as reported on Fig. 12.

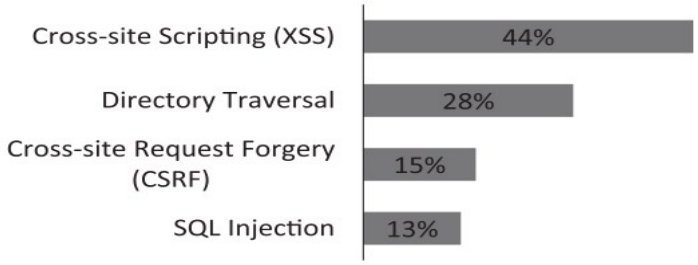

Fig. 12. Different cyber-attacks blocked by secure cloud hosting

4.2.4.4.3 Cross-Site Request Forgery (CSRF) and Command Injections

Specific malware injection attacks can also occur like *command injections* and *CSRF. Cross-Site Request Forgery* (CSRF) is an attack that forces an end user to execute unwanted actions on a web application in which they are currently authenticated. It is, actually, an attack that tricks a user into performing unwanted actions on a web application and exploits the trust a web application has in the user's browser. The victim logs into a web application (e.g., an online banking site) and receives a session cookie, which the browser automatically includes in subsequent requests to the application. Then, the attacker creates a malicious request that performs an action on the target application (e.g., transferring money, changing account settings). The attacker tricks the victim into visiting a malicious website or clicking a link that triggers the crafted request. Following, the victim's browser sends the malicious request to the target application, including the session cookie and The application processes the request as legitimate because it includes the victim's valid session cookie. Finally, the application performs the action (e.g., transferring money) without the victim's knowledge or consent.

The *command injection* is actually, a type of attack where an attacker exploits a vulnerability in a web application to execute arbitrary commands on the host operating system. This occurs when user input is improperly sanitized and passed to a system shell for execution. A web application takes user input (e.g., from a form, URL parameter, or API request) and uses it to construct a system

command. The attacker, then, provides malicious input that includes additional commands or special characters to manipulate the system command and the application passes the malicious input to a system shell without proper validation or sanitization. Finally, the attacker's injected command is executed with the same privileges as the web application, potentially causing significant damage.

4.2.4.5 Container escape attacks

A *container* in cloud computing is a lightweight, portable, and self-sufficient software package that includes everything needed to run an application, such as code, runtime, system tools, libraries, and dependencies. Containers allow applications to run consistently across different computing environments, whether on-premises, in the cloud, or in hybrid environments.

Containers have become a cornerstone of modern cloud computing, enabling developers to create, deploy, and manage applications with increased efficiency and portability. However, while containers offer many benefits, they also introduce unique security challenges and vulnerabilities that can significantly impact cloud privacy. Organizations leveraging containerized applications must understand these vulnerabilities and proactively implement strategies to safeguard sensitive data and maintain privacy in the cloud.

Container Escape Attacks: A container escape attack occurs when an attacker exploits a flaw in the container runtime or kernel to break out of the container’s isolated environment. By exploiting vulnerabilities in the container engine or the host operating system, attackers can gain access to the host system, potentially accessing other containers running on the same host. This is a serious threat as it can compromise the entire cloud infrastructure, enabling attackers to access not only the data within the container but also the underlying system, resulting in large-scale data breaches.

4.2.4.6 Cookie Poisoning Attacks

In this kind of attack, the contents of a cookie are changed to get access to an unauthorized application or web page. The cookie contains sensitive credentials about the user’s data (could be done via an XSS attack). So, when the hacker gains access to a cookie, then he/she also gains access to the content within the cookie and can perform illegal activities (e.g., if the cookie includes a session key, he/she can use it to perform actions in the server on behalf of the victim). These attacks are also considered as multilayer attacks (affects also the IaaS). Moreover, these attacks can be considered that affects primary the *Presentation layer*[55] in Cloud Computing (and secondary the *Application layer*[56] in Cloud Computing).

[55] The Presentation Layer in cloud computing refers to the layer responsible for managing the formatting, presentation, and display of data to end-users or systems. This layer acts as an intermediary between the Application Layer (where the core logic and functionalities reside) and the User Interface (UI) that interacts with end-users. It ensures that the data provided by cloud applications is translated into a user-friendly format and is displayed in a way that users can easily understand and interact with.

[56] The Application Layer in cloud computing refers to the top layer of the cloud architecture that directly interacts with end-users and provides cloud-based applications and services.

4.2.5 Cloud privacy attacks on the PaaS cloud layer

This delivery model contains flaws that might lead to attacks on PaaS services, such as, password reset attacks, man-in-the-middle attacks, cloud malware-injection attack and phishing attacks. These attacks are analysed below in respective sub-sections.

4.2.5.1 Password-related attacks

A password attack definition encompasses various methods malicious actors attempt to breach a system or account by compromising user credentials. These intrusions often target weaknesses in authentication systems or take advantage of easily predictable or commonly used passwords. We consider these attacks more related to PaaS layer, as PaaS vendors typically provide encryption for data in transit by default or provide mechanisms to enable it. In PaaS environments, data in transit (password protection during transit over a TCP socket) has a higher security priority than data at rest, because PaaS has complex workflows and tends to integrate with multiple external systems. Nevertheless, passwords can be also leaked at the other two (SaaS, IaaS) cloud layers.

In other words, it is the process in which the attackers try every possible character combination to recover the password from encrypted data that they have intercepted (e.g., via a MiTM attack). The password can be recovered to its own original content by spending extra effort through using high computational resources, tools, and techniques. However, purchasing high computing supercomputers to decrypt encrypted information is not a wise decision which may work for specific tasks and is too costly in terms of financial burden. However, with CC, end-users have access to very powerful machines to execute their high computational task which is far better than the machines operating in the traditional infrastructure. So, it could be effective for some activities (like accessing user credentials), buying powerful supercomputers to decrypt encrypted data, but at the same time, might be considered as not a smart move and comes with a hefty financial cost. CC, on the other hand, gives end users access to extremely powerful computers that can do high-computational tasks much more effectively than the machines found in traditional infrastructure. Nevertheless, there are various forms of password-related attacks:

- A *Brute Force attack* is a method used by hackers to crack passwords, encryption keys, or other login credentials by systematically trying all possible combinations. The attacker uses software that rapidly guesses passwords until it finds the correct one, often targeting weak or commonly used passwords. The primary goal is to gain unauthorized access to sensitive data or systems.
- *Dictionary-based attack*, is a type of brute force attack where hackers try to guess a user’s password to their online accounts by quickly running through a list of commonly used words, phrases, and number combinations. When a dictionary attack has successfully cracked a password, the hacker can then use this to gain access to things like bank accounts, social media profiles, and even password-protected files. This is when it can become a real problem for the attacker’s victim. Although the material is encrypted, this does not exclude the encryption breakage. With extra work and the use of powerful tools, methodologies, and computer resources, the password can be restored to its original content.

- *Password spraying.* It involves trying a large number of common passwords which are checked over a set of accounts and not just one account. Attackers go to great lengths to avoid detection during password spraying. Usually, they will conduct reconnaissance first to limit the number of logins attempts and prevent account lockup.
- *Credential harvesting* (see above section 4.2.4.2 Authentication attack)
- *Keylogger attack.* A keylogger is spyware that records a user’s activity by logging keyboard strokes. Cybercriminals use keyloggers for stealing a variety of sensitive data, from passwords to credit card numbers. In a password attack, the keylogger (probably injected somehow into the user system) records not only the username and password but also the website or application where those credentials are used, along with other sensitive information.

4.2.5.2 Man-in-the-middle attack

A Man-in-the-Middle (MiTM) attack occurs when an attacker intercepts and manipulates the communication between two parties without their knowledge. In cloud environments, MiTM attacks target confidentiality, integrity, and authenticity of data, posing significant risks to cloud privacy. A Man-in-the-Middle (MiTM) attack is not limited to cloud environments. It can occur in any *networked environment*, including *local networks*, public Wi-Fi, corporate intranets, and cloud services. It can be considered an attack at the Platform-as-a-Service (PaaS) cloud layer because it targets the communication between users and the cloud-based applications or services hosted on the PaaS platform. PaaS environments heavily rely on APIs for communication between applications, services, and users. A MiTM attacker can intercept unencrypted API traffic, stealing sensitive data (e.g., credentials, tokens) or injecting malicious payloads.

MiTM attacks compromise cloud privacy in several ways, regarding: Data Breaches for sensitive information, such as credentials, personal data, or business records, Reputation Damage because breaches resulting from privacy violations erode trust in cloud providers and Regulatory Fines because non-compliance with privacy laws (e.g., GDPR, CCPA) due to data breaches can lead to financial penalties.

This would be one kind of eavesdropping attack that involves interception between a couple of parties. Here, the attackers try to interfere between two communicating entities, where one is a customer and the other is the (cloud) authentication sub-system, a service or application. They insert themselves right into the middle of the communication path and try to impersonate one of the two communicating parties. From this, he may intercept or modify the communication. Normally, network traffic travels between two computers that communicate with each other over the Internet (this is also common in cloud environments where the Internet is used to access cloud services). In such an attack, while transmitting messages between two computers, the attacker intercepts messages in a public key exchange and then retransmits this message by substituting the current with their own public key using a method where the original users still appear to be communicating with each other. So, he can trick one or both parties such that they are utilizing his secret key(s) for communication.

More specifically, we have cases in the context of authentication/identity verification:

a. There is no encryption and identity verification so the attacker can easily impersonate one of the two parties

b. There is encryption but no identity verification. So, the attacker can intervene in the public key exchange phase to impose its own secret key to both parties.

c. There is encryption and identity verification. This is the hardest case for the attacker. He/she needs to impersonate digital certificates, which is more difficult but still possible (e.g., he/she can create and publish a certificate to one of the parties that claims to map to the other party. If certificates are not verified, then this could do the trick. Otherwise, CAs need to be infiltered so as to produce verifiable certificates).

There are multiple alternative versions of this attack:

a. the intermediary can intercept the communicated data, especially when they are sent in a non-encrypted form

b. he/she can modify the data which is even worse

c. he/she can perform actions on behalf of the customer which is very dangerous

The man-in-the-middle (MiTM) attack process has a two-stage approach: interception and decryption.

Interception

During the interception step, the cybercriminal attempts to put themselves between the client and server—typically a client/user and web application. Depending on the type of man-in-the-middle attack, there are a few ways the attacker could approach this:

- Creating a non-secure Wi-Fi network or rogue hotspot in a crowded area for people to connect such that the attacker can view their information.
- Accessing a (public) Wi-Fi network, typically by taking advantage of a weak password or by installing a packet sniffer to analyze traffic and scan for vulnerabilities, points of entry, and ideal targets.
- Creating a fake website and route the user through phishing links in emails or redirecting them from the HTTPS site (they intended to visit) through DNS spoofing. The spoofed DNS means that the DNS entry (in the respective DNS server advised by the user browser) is poisoned and redirects users from a legitimate to a malicious web site.

Decryption

After targets are determined and fall for bait, cybercriminals use data capture tools to transmit any login information and web activity back to them and decrypt it into readable text if they come in a encrypted form. During the decryption phase, the intercepted data becomes usable to the criminal.

Some types of Man-in-the Middle (MITM) Attacks include Wi-Fi eavesdropping, DNS spoofing, HTTPS spoofing, Secure Sockets Layer (SSL) hijacking, ARP poisoning.

- *DNS spoofing*: It is also known as DNS cache poisoning, involves corrupting the DNS (Domain Name System) query process. DNS spoofing can be considered a specific technique that might be used in the broader context of a MitM attack. By

manipulating DNS responses, an attacker can redirect a victim’s traffic through a device or server controlled by them, effectively placing themselves "in the middle" of communication. Once the traffic is redirected, the attacker can monitor, collect, or alter the data transmitted between the victim and what they believe is a legitimate site.

- *Wi-Fi eavesdropping:* Attackers can eavesdrop on Wi-Fi networks by attaching packet sniffers to insecure Wi-Fi hotspots. The insecure Wi-Fi network may be a public or private network with weak passwords that the attacker cracks using password cracking software (e.g., John the Ripper). Alternatively, it may be a Wi-Fi network owned by the malicious actor and disguised as a free, public network. The moment unsuspecting victims connect to insecure hotspots, the attacker becomes an MitM, able to access all their communications.
- *ARP spoofing,* also known as ARP poisoning, is another technique that can facilitate Man-in-the-Middle (MitM) attacks. ARP spoofing is a direct method to perform MitM attacks within a local network. Once the attacker has successfully associated their MAC address with the IP address of a target device (e.g., a gateway or server), they can intercept, modify, or redirect the traffic between two devices without their knowledge.
- *SSL hijacking:* In SSL hijacking, the cybercriminal intercepts SSL/TLS network traffic and sends fake SSL certificates to both the client and server during a TCP handshake. This allows the attacker to impersonate the server, force victims to connect to unsecured websites, and control the session.
- *HTTPS spoofing:* Also known as homograph attack, Hypertext Transfer Protocol Secure (HTTPS) spoofing starts with a victim who requests to securely connect to a website. In response, a hacker sends the victim a hoax certificate that belongs to a lookalike but malicious version of the target website. The victim’s browser verifies the certificate, believing it to be a trusted site, which then allows the cyber attacker to decrypt the communication. By tricking victims into trusting a fraudulent SSL certificate, attackers can steal sensitive data, such as login credentials, financial details, and session tokens. Users think that they are in a secure HTTPS website, but they are interacting with a fake or malicious website.

Spoofing attacks: In general, Spoofing is a specific type of cyber-attack in which someone attempts to impersonate a computer, device, or network to masquerade as a legitimate entity for malicious purposes. An IP spoofing attack may replace the source IP address in a network packet so as to seem to come from an existing device. So, we have an impersonation of that device such that the packet seems to come from that device and not by the attacker. IP spoofing could be exploited in the context of a DDoS attack for the masking of the botnet devices involved in such an attack.

Similarly, a DNS spoofing attack may cause a penetrated DNS server to return wrong DNS responses including incorrect IP addresses thereby redirecting network traffic to an attacker's

system. A virtual network may be a victim of ARP spoofing[57] thereby causing an attacker VM to access packets of other VMs. ARP affects routing so packets end up in the wrong place. As such, the attacker may have access to such a packet. These attacks can also be considered as multilayer attacks (affects primary the PaaS and secondary the IaaS). Indeed, they primarily target PaaS because they exploit communication between users and cloud applications, intercepting API traffic or redirecting requests. They, secondarily, affect IaaS by targeting network-level vulnerabilities, such as virtual machine communication or DNS resolution in cloud infrastructure. Essentially, exploits PaaS layers, but affects also IaaS layers.

4.2.5.3 Cloud malware-injection attack (PaaS layer case)

A Cloud malware-injection attack is a type of cyber-attack that involves injecting malicious software, such as viruses or ransomware[58], into cloud computing resources or infrastructure, but mainly in data and applications, application runtime, containers, serverless functions and databases (can also infect IaaS and SaaS layers). This can allow the attacker to compromise the affected resources and steal or destroy data, or to use the resources for their own purposes. A malicious code may be inserted in an application (also via XSS) if the cloud application has an insecure interface/API that exhibits vulnerabilities ("insecure APIs", "human error" in terms of utilizing an insecure interface or "zero-day"*)* (Ko et al., 2013; Owens, 2010). In this layer, the attacker exploits vulnerabilities in a platform-hosted application to inject malicious code. For instance, an attacker could exploit a misconfigured API or insecure file upload feature in a PaaS-based web app to inject a malicious script. This script could then steal sensitive data, disrupt services, or spread ransomware, compromising the entire application and its users. In this way, we have an attacker who receives files in insecure channels and modifies them to inject malicious code into them. The transmission of malicious code may then be ignored by security systems for which it seems as if a normal file is being sent.

This attack is also considered as a multilayer attack (affects also SaaS and IaaS[59]).

There are several ways in which attackers can inject malware into cloud resources, including:

- Exploiting vulnerabilities in the cloud infrastructure or in the systems and applications running on the cloud.

---

[57] An *ARP (Application Resolution Protocol) spoofing*, also known as ARP poisoning, is a type of Man in the Middle (MitM) attack that allows attackers to intercept communication between network devices. The attack works as follows:
The attackers must have access to the network. They scan the network to determine the IP addresses of at least two devices—let's say these are a workstation and a router. The attackers use a spoofing tool, such as Arpspoof or Driftnet, to send out forged ARP responses. The forged responses advertise that the correct MAC address for both IP addresses, belonging to the router and workstation, is the attacker's MAC address. This fools both router and workstation to connect to the attacker's machine, instead of to each other. The two devices update their ARP cache entries and from that point onwards, communicate with the attacker instead of directly with each other. The attackers is now secretly in the middle of all communications.

[58] Ransomware attacks are where attackers encrypt an organization's files or systems and demand payment for decryption. Unlike privacy breaches, ransomware *focuses on system disruption rather than stealing sensitive data*. This type of attack has affected hospitals, manufacturing companies, and even government agencies, crippling operations until a ransom is paid.

[59] In IaaS environments this attack targets Virtual Machines (VMs), Storage, Hypervisors and Network. One method is the "Compromised VM Images" where attackers inject malware into shared VM templates/images, spreading infected instances. In SaaS environments this attack targets Web Applications, Cloud-Based Collaboration Tools and User Data. One method is the "Cloud Ransomware" where infecting cloud storage (e.g., Google Drive, OneDrive, AWS S3) to encrypt user files.

- Adding a malicious service module to a SaaS or PaaS system, or an infected VM to an IaaS system, and diverting user traffic to it.
- Gaining unauthorized access to cloud accounts and injecting malware using malware-infected files.

4.2.5.4 Phishing attacks/Social engineering attack

Phishing attacks usually consist of phony emails that lead to a fake/malicious website that the user could visit and potentially be deceived by it to unveil personal data (like credentials for the legitimate web site that is mimicked by the fake one). Both consumers and enterprises are impacted by these threats in also a PaaS cloud environment. By employing phony websites, emails, or blogs, these attackers persuade the victims to divulge their most personal information, such as their password, by using many deceiving methods that focus on psychological manipulation.

Phishing and social engineering attacks pose a significant threat to PaaS environments by exploiting human vulnerabilities to bypass technical defenses. Organizations must prioritize user education, implement robust access controls, and adopt a proactive security posture to mitigate these risks and protect their cloud-based applications and data.

The Phishing/Social Engineering in PaaS, target Developers and Administrators. They send phishing emails or messages disguised as legitimate communications (to trick developers or admins into revealing credentials, API keys, or access tokens. They exploit Trust in Third-Party Integrations, because PaaS platforms often integrate with third-party tools and services. Attackers impersonate these tools to trick users into granting access or installing malicious plugins. They compromise User Accounts, once credentials are obtained, attackers can log into the PaaS platform, manipulate applications, steal data, or inject malicious code. At the same time, they perform API Key Theft, because phishing attacks can trick users into revealing API keys, which attackers use to interact with PaaS services programmatically, exfiltrating data or disrupting operations.

The impact of attacks on PaaS is related to Data Breaches, Service Disruption, Reputational Damage and Financial Loss.

Phishing assaults are a usual problem for internet users. For instance, phishing attempts frequently target Facebook users. (Li et al., 2011[60]). These attacks are also considered as multilayer attacks (affects also the SaaS and the IaaS[61]). Social engineering attacks rely on human psychology rather than technical vulnerabilities. One common method is Business Email Compromise (BEC), where attackers impersonate executives or trusted contacts to manipulate employees into transferring funds or disclosing confidential information.

[60] Li, T., et al., 2011. “*LARX: Large-scale anti-phishing by retrospective data-exploring based on a cloud computing platform*”. In: Proceedings of the Computer Communications and Networks (ICCCN), 2011 Proceedings of 20th International Conference on, 2011. IEEE

[61] In IaaS environments this attack targets virtualized computing resources over the internet, such as virtual machines, storage, and networking. Attackers target IaaS environments to gain control over infrastructure components, with methods as the “Credential Theft” and “Impersonate” (attackers impersonate IT administrators or cloud providers to trick users into granting access or revealing sensitive information).

Moreover, phishing attacks are typically fraudulent email messages which are directed to spoofed websites. In PaaS cloud environment, these attacks affect both enterprise and users. This is a type of social engineering attack.

We could also have malicious attachments involved in emails or messages that attempt to convince the user to transfer money (depending on the kind of phishing attack involved). Internet users suffer from phishing attacks. Famous web browsers provide plug-ins (or email clients can provide plug-ins) to defend against these attacks, but a complete client-side solution is not being applied so far.

4.2.6 Cloud privacy attacks on the IaaS cloud layer

At this layer, the first thing attackers do is control the operations of hosted virtual machines (VM), opening the possibility of an attack on other hosted VMs or on the hypervisor. Virtual machine (VM) introspection is another powerful technique for determining the specific aspects of guest VM execution from outside the VM. Unfortunately, existing introspection solutions share a common questionable assumption. This assumption is embodied in the expectation that original kernel data structures are respected by the untrusted guest and thus can be directly used to bridge the well-known semantic gap. An attack called DKSM can effectively foil existing VM introspection solutions into providing false information.

4.2.6.1 DKSM (Direct Kernel Structure Manipulation) attacks

These attacks are particularly dangerous in IaaS environments, where large amounts of user data are stored in shared infrastructure and where multiple virtual machines (VMs) share the same physical hardware. Direct Kernel Structure Manipulation (DKSM) attack is a sophisticated rootkit-based attack that, essentially, targets the kernel of an operating system in cloud infrastructure. DKSM attacks modify key kernel structures to gain persistent, stealthy control over cloud workloads. DKSM (Direct Kernel Structure Manipulation) attack is a technique to bypass isolation mechanisms (memory isolation, user/kernel space separation, etc) and exploits the semantic gap[62]. The semantic gap often plays a crucial role in the effectiveness of DKSM attacks because it can allow attackers to exploit mismatches between a system's intended behavior and its actual implementation. By implementing strong encryption, isolated key management, and access control policies, cloud providers can mitigate the risks of DKSM attacks.

4.2.6.2 Virtual machine (VM) introspection

Virtual Machine Introspection (VMI) attacks exploit the ability to monitor and manipulate virtual machines (VMs) from a privileged position outside the guest operating system. In IaaS (Infrastructure as a Service) environments, where multiple VMs run on shared hypervisors, a

[62] A key challenge here, is the so-called *semantic gap* between the external and internal observations of a VM. Specifically, from outside the VM, we can get a view of the VM at the virtual machine monitor (VMM) level, which includes its register values, memory pages, disk blocks and low-level events, whereas from inside the VM, we can observe semantic-level entities and events. This semantic gap is formed by the vast difference between external and internal observations, manifesting the main challenge for VM introspection. Semantic gap can result in misalignments between the security policies and the enforcement mechanisms.

malicious hypervisor, administrator, or compromised management layer can leverage VMI to perform attacks, such as data exfiltration, process manipulation, and privilege escalation. These attacks exploit the monitoring capabilities provided by hypervisors (like VMware, Xen, or KVM), which are typically used for debugging, security monitoring, or virtualization management.

Types of these attacks are Hypervisor attack (Escape Attack and Privilege escalation[63]), Introspection data[64] and Memory Injection and Modification[65].

4.2.6.3 Side channel attacks

An often-overlooked side-channel attack is the one concerning the level 3 cache which is sometimes shared between different co-located VMs (by the same host). A side channel attack is defined as a method of stealing information indirectly by exploiting unintended information leakage, such as side-channel signals generated during computation. These attacks can compromise cryptography systems by extracting secret keys or plaintext data used in encryption schemes. This attack is a kind of the Cross-VM attacks category.

Through the side channel attack, an attacker sharing the same cache as the victim can monitor the cache access behaviour of the victim. For example, the attacker is able to monitor cache timing information by measuring the execution of different operations on the victim's VM (Zhang and Lee 2014). If this cache incorporates unencrypted data, then it is possible that an attacker having a hosted VM directly accesses data in this cache placed by another co-hosted VM of a legitimate client. However, if the data are encrypted, then encryption should be somehow broken such that the attacker could get into his/her hands the original data.

Based on the above analysis, there are two main steps involved inside channel attacks:

(i) *Placement*: the malicious VM (of the attacker) resides on the same physical machine and

(ii) *Extraction*: the malicious VM directly accesses, due to improper isolation, the L3 cache which is shared with other VMs on the same physical machine. So, it can directly access data that is placed in the L3 cache by these VMs.

As such, cache side channel is a Cross VM attack that has as its source two main causes: (a) multi-tenancy (due to cloud’s nature) and (b) the improper resource isolation (a certain vulnerability). (Xiao and Xiao, 2013[66]).

Another major kind of side channel attack due to the sharing of cloud infrastructure is the time-based side-channel attack. In such an attack, timing channels can be exited at any point in time and malicious users co-located on the same infrastructure can potentially steal sensitive data from

[63] An *Escape Attack* occurs when an attacker breaks out of an isolated environment, like a container or virtual machine, to access the host system or other resources. *Privilege Escalation* involves exploiting vulnerabilities to gain higher-level permissions, such as root access, enabling unauthorized control over systems, data, or applications. Both compromise cloud security at IaaS cloud layer.

[64] An *Introspection Data Attack* in the IaaS cloud layer involves exploiting hypervisor-level introspection tools, which monitor virtual machines (VMs), to access sensitive data or metadata. Attackers can abuse these tools to extract information like encryption keys, credentials, or VM configurations, compromising the security and privacy of cloud-hosted workloads.

[65] A *Memory Injection and Modification Attack* in the IaaS cloud layer involves injecting malicious code or altering the memory of virtual machines (VMs) to manipulate processes, steal data, or escalate privileges. Attackers exploit vulnerabilities in shared resources or hypervisors to compromise VM integrity, bypassing traditional security measures and gaining unauthorized access.

[66] Xiao, Z., Xiao, Y., 2013. “*Security and privacy in cloud computing*”. Commun. Surv. Tutor. IEEE 15 (2), 843–859.

other legitimate users. A timing attack enables the attacker to extract secret information by analyzing the time taken to respond to different queries. As the state of cache memory directly depends on its memory access timing. If a time-based side-channel attack concerns the cache, then it coincides with the cache side-channel attack that exploits timing information as also explained. Different types of processes take different execution times; that time difference could be used for breaking cryptographic keys like AES, RSA, or ElGamal keys. In this sense, if we combine cache & time-based attacks, encrypted data can be illegally decrypted in their original form.

Side channel attacks are, probably, one of the most challenging threats to address (unless you have proper isolation). Better resource isolation is the right measure to follow in this case. Other mitigation techniques for side-channel attacks includes apart from the Isolation (which use dedicated hardware or cores for sensitive tasks), Constant-Time Algorithms (implement cryptographic operations resistant to timing analysis), Noise Injection (add randomness to obscure data patterns), Access Control (restrict physical and logical access to resources) and Regular Patching (update systems to address vulnerabilities).

4.2.6.4 VM rollback attack

In this attack, the attacker, once able to penetrate the hypervisor or have unauthorized access to the administration console (by compromising the admin account or by exploiting weak APIs), can request to roll back a VM that the hypervisor controls to a known vulnerable state. Then, it is much easier for the attacker to penetrate that VM, stealing sensitive data like encryption/secret keys.

In a VM rollback attack, a compromised hypervisor runs a VM from an old snapshot without the user’s awareness. Here, a user is the one who uses cloud service and is the owner of guest VM(s). Since a part of the history of VM’s execution is lost, an attacker can bypass some security checks in the VM or undo some security critical updates. For example, an attacker launches a brute-force attack to guess the login password of a VM.

VMs roll back to their previous state if an error occurs. Unfortunately, this factor can re-expose them to security vulnerabilities that can lead to data privacy and integrity issues and therefore attackers will exploit the compromised hypervisor.

In old IaaS platforms, it is difficult to identify the normal resume/suspend and rollback operations in old systems (malicious ones) so as a result they are vulnerable to rollback attacks. Before, systems simply disabled all these features or required continuous feedback from customers to defend against these attacks. However, if a hypervisor is vulnerable, then not only this but also other attacks are possible. As such, stopping just one attack might not be enough.

To balance the security and functionality Xia et al. (2015) proposed a method where users securely log all suspend/ resume and rollback operation inside the trusted computing domain.

4.2.6.5 Stepping-stone attack

The attackers do not attack from their own machine in a stepping-stone attack but rather from intermediary hosts by using previously compromised VMs. The attacker can reside of course outside the cloud and connect to the compromised VM to perform the attacks. In this way, VMs are used to

conduct a range of attacks towards the main target (e.g., DDoS, MiTM) without disclosing the true identity of the attacker. In this scenario, the service providers in the IaaS layer are attackers and victims at the same time. Attackers as VMs from their cloud conduct attacks. Victims as VMs in their cloud are compromised and/or receive an attack. In any case, the IaaS providers must take the necessary security measures to prevent such attacks or at least mitigate their impact (e.g., through VPNs and network partitions the infection cannot be spread across the whole cloud infrastructure).

4.2.6.6 VM escape (privilege escalation attack)

In the cloud virtualization model, there is normally a privileged VM and several unprivileged VMs. A privilege escalation attack is to escalate the privileges of an unprivileged VM. A privileged VM offers an administration interface to the hypervisor, so making a VM privileged means that the VM can access the hypervisor and thus the VMs that it controls. In this mode, the Boolean value in the hypervisor controls whether the VM is privileged or not. After escalating the privilege of a compromised VM the attacker can harm the VM and potentially other co-hosted VMs. This could be one way to conduct VM escape but usually VM escape means breaking the isolation level and there could be different ways to achieve that. In addition, privileged VMs might not exist in all possible cases.

If the attackers escape the resource/VM isolation level, they may compromise the hypervisor and as a result may control the entire guest OS. If the attacker controls the hypervisor, it can also do anything to the VMs on the host system (You et al., 2012[67]).

In other words, a virtual machine (VM) escape is an exploit in which an attacker runs code on a VM that lets the operating system (OS) running within it break out and interact directly with the hypervisor. By exploiting a vulnerability in the hypervisor's code, guest OS or applications running on the VM, the attacker can directly interface with and compromise the underlying physical resources and VMs.

The exploitation of the hypervisor is called a hypervisor-level VM escape attack. Another way to execute a VM escape attack is to exploit the vulnerabilities in VM applications. Actually, exploiting vulnerabilities in a VM escape attack involves identifying weaknesses in the hypervisor or VM isolation mechanisms. Attackers use these flaws to break out of the VM's confined environment, gaining unauthorized access to the host system or other VMs. This compromises the entire cloud infrastructure, exposing sensitive data and resources. Either way, a successful attack lets the attacker escape the isolation of a VM and then access and control the host OS and all the other VMs running on it. The attacker might also be able to access any sensitive information stored on the physical machine and even on the network to which it's connected.

4.2.6.7 Attacks from malicious insiders

Mainly, a malicious insider is an employee or a third party, who sometimes has administrative or sufficient rights or substantial privileges and thus the power to do harm to a cloud environment. Maybe for the insider it could be the case that he/she can configure or manage some cloud services

[67] You, P., et al., 2012. “*Security issues and solutions in cloud computing*”. In: Proceedings of the Distributed Computing Systems Workshops (ICDCSW), 2012 32nd International Conference on. 2012. IEEE.

so he/she could have increased rights for those services. The same might hold for a third party that might manage cloud resources on behalf of the cloud provider (although a third-party is usually trusted). This kind of attack occurs in the cloud when the person, employee or staff who know how the system runs, can implant malicious codes to destroy everything in the cloud system, or to disclose sensitive data or sell them in the black market / dark web, can use resources for his/her own purposes degrading their performance, can launch attacks on other systems (like DDoS ones) and so on.

Malicious Clients, Malicious Cloud Provider/Broker etc. are all the other terms which can also be used as an alternative to malicious insiders.

The CC industry is aware of the threat of insider attacks (Fig. 13). This threat is much more dangerous in cloud environments because it may gain a lot of important information from the cloud data. It is also very important to know the scope of the insider attacks and their effect to better defend against such threats. We need to understand whether insiders exposed important information from both actors and attack surface. This could be considered as a general strategy to resolve such an attack and then examine if such strategy is followed and to what degree (Duncan et al., 2012[68]).

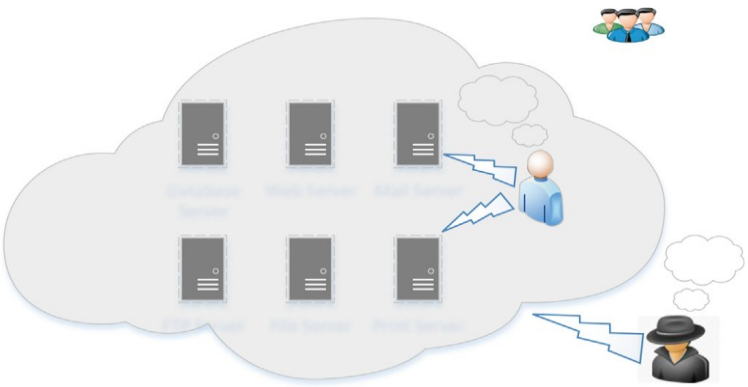

Fig. 13. An overview of internal and external attacks

4.2.6.8 Cross-VM attacks

Cross-VM attacks refer to a category of attacks where an adversary targets multiple virtual machines (VMs) running on the same hypervisor, aiming to break the isolation between them. These attacks can occur when one VM, usually compromised by an attacker, or other compromised VMs hosted on the same physical machine, are used by different customers. Cross-VM attacks exploit, essentially, the very nature of multi-tenancy. Multi-tenancy enables different customers to have respective VMs that can co-reside on the same physical machine. VMs running on the same physical server share the same resources. Hence, cross-VM attacks are possible to take place in the multi-tenant virtualized environment. Most researchers focus on cross-VM attacks which primarily target the cache memory (see Side Channel attack above). Subcategories of cross-VM attacks include: Side-Channel Attacks (exploit shared hardware (e.g., CPU caches) to extract sensitive data), Co-

[68] Duncan, A.J., Creese, S., Goldsmith, M., 2012. “*Insider attacks in cloud computing*”. In: Proceedings of the Trust, Security and Privacy in Computing and Communications (TrustCom), 2012 IEEE 11th International Conference on. 2012. IEEE.

Residency Exploits (target VMs on the same physical host to access shared resources), Network-Based Attacks (intercept or manipulate traffic between VMs) and Hypervisor Exploits (abuse hypervisor vulnerabilities to compromise VM isolation).

To protect this scenario (in the IaaS cloud layer) it should be provided an active response against these types of attacks. It is very difficult to protect against types of attacks through firewalls because firewalls usually protect the information in network packets and filter incoming information (Kourai et al., 2012[69]).

4.2.6.9 Return-oriented programming attack (ROP)

In a virtualized environment the code reuse (or 'return-into-library') is called VM Programming (Wang and Jiang, 2010[70]). In this type of attack (the '*code reuse*' essentially), the stack is overwritten with addresses that are pointing to a function call stack in the standard library (stack overwriting technique). A typical ROP exploit includes illegitimate manipulation of a function call stack used by a thread of a process; the illegitimate manipulation intended to alter the original functionality of the respective process. For instance, an exemplary ROP exploit may manipulate the function call stack to force the host system to execute only a subset of instructions of the original process, and/or to execute such instructions in a sequence, which differs from the sequence of instructions of the original process (Ding et al., 2012). The attackers attempt to control the guest OS to access the hypervisor or to penetrate the functionalities of another guest OS and underlying host OS by taking over its functionality.

Return-oriented programming provides a fully functional "language" that an attacker can use to make a compromised machine perform any operation desired. In a 'return-into-library' attack, an attacker hijacks program control flow by exploiting a *buffer overrun vulnerability*[71]. Instead of attempting to write an attack payload onto the stack, the attacker instead chooses an available library function and overwrites the return address with its entry location. Further stack locations are then overwritten, obeying applicable calling conventions, to carefully pass the proper parameters to the function so it performs functionality useful to the attacker.

4.2.6.10 Reused IP address attacks

Each node of a network has an IP address which is allocated to a particular user. When that user leaves the network, the IP address associated with him/her can be assigned to a new user. The chances of accessing previous user data by the new user exist as the address still exists in DNS cache

[69] Kourai, K., Azumi, T., Chiba, S., 2012. "*A self-protection mechanism against stepping-stone attacks, for IaaS clouds*". In: Proceedings of the Ubiquitous Intelligence & Computing and 9th International Conference on Autonomic & Trusted Computing (UIC/ATC), 2012 9th International Conference on. 2012. IEEE.

[70] Wang, Z., Jiang, X., 2010. "*Hypersafe: A lightweight approach to provide lifetime hypervisor control-flow integrity*". In: Proceedings of the Security and Privacy (SP), 2010 IEEE Symposium on. 2010. IEEE.

[71] Buffer overruns and overflows problems can affect all types of software. They typically result from malformed inputs or failure to allocate enough space for the buffer. If the transaction overwrites executable code, it can cause the program to behave unpredictably and generate incorrect results, memory access errors, or crashes.

and hence the data belonging to one person can be accessed by another[72] (Amara et. al., 2017). Of course, a node can be accessed by multiple users at the same time due to multi-tenancy. For the case of VMs, then one tenant maps to one VM IP address. VM access is done based on SSH and specific keys which are differentiated between the tenants. Even if an IP address maps to a previous VM, the VM might have been destroyed or stopped (so it is inaccessible) and the access to the VM cannot be done with the new tenant credentials. However, if the original IP address had trust relationships, DNS entries, or firewall rules associated with it, the attacker can take advantage of these preexisting configurations. For example:

- Exploiting trust in whitelisted IPs.
- Receiving traffic intended for the original owner.
- Misusing credentials tied to the old IP address.

These attacks are also considered as multilayer attacks (affects also PaaS). They are related to the PaaS layer because PaaS applications often rely on dynamic IP allocation and shared resources. If an attacker gains access to a previously used IP address, they can intercept traffic, spoof services, or exploit residual configurations, compromising the security of PaaS-hosted applications and data.

Moreover, these attacks can be considered that affects the *Network layer*[73] in Cloud Computing (and secondary the *Application layer*[74] in Cloud Computing). The network supports intermachine communication as well as external access/communication to the machines from the outside. An intruder may attack a cloud system through its network. This may in turn deteriorate the quality of the cloud services (e.g., by introducing unnecessary network delays) and may even put data privacy/confidentiality at risk (if data are not encrypted at transit). For our analysis, we consider three types of network-based attacks as elaborated below. The Port Scanning is a preparatory step

[72] Amara, N., Zhiqui, H., Ali, A., "*Cloud Computing Security Threats and Attacks with their Mitigation Techniques*", 2017 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery, 978-1-5386-2209-4/17© 2017 IEEEDOI 10.1109/CyberC.2017.37244 (pp. 249)

[73] In cloud computing, the Network Layer refers to the part of the cloud architecture responsible for managing the network infrastructure, which facilitates communication between different components of the cloud environment (such as virtual machines, storage, applications, and users). The network layer handles the routing, switching, and management of traffic across different resources, ensuring that data can be sent and received securely and efficiently.

[74] The Application Layer in cloud computing refers to the top layer of the cloud architecture that directly interacts with end-users and provides cloud-based applications and services. The applications running on a cloud may be exposed to various attacks, for example by injecting code which may trace execution paths and exploit this information for malicious purposes. Similarly, the protocols (HTTP, DNS, VoIP, SMTP) implemented to provide services on a cloud system are vulnerable to attacks and any running application may unintentionally use them as a source of intrusion. Moreover, in a cloud system, the architectural components being shared (as an application in PaaS) may be exploited by an attack performing malicious activities.

for an attack like DoS (Riquet et al., 2012[75], Scarfone and Mell, 2007[76]) and a Botnet[77] is the medium via which such an attack can take place (McMillan, 2009[78]).

4.2.6.11 Sniffing Attacks

The attacker launches these attacks by means of programs that allow a host to capture flowing packets in an Ethernet network by means of inserting the MAC address of the host's Network Interface Card (NIC) into the malicious code. If the information in these packets which is being transferred is not encrypted, it can be compromised by these programs[79]. These attacks are also considered as multilayer attacks (affects secondary PaaS and SaaS) but can be also categorized in the *Network layer* in Cloud Computing.

4.2.6.12 VM image supply attacks

A malicious code can be placed inside a VM image which is then replicated during the creation of a VM. In this case, a VM image seems to come from a legitimate source but contains vulnerable software. With proper configuration of the VM management system, it can detect the malicious code and stop the VM image creation / updating process and therefore it is difficult for malicious software to penetrate beyond the individual VM session (but not impossible). In particular, this type of penetration from malicious software can cause significant problems by infecting the underlying host, the hypervisor and beyond through the launching of additional attacks like the VM escape one.

4.2.6.13 VM scheduler-based attacks

A few vulnerabilities of VM scheduler (lack of visibility, shadow IT) may result in resource stealing or theft-of-service. For example, a VM may be scheduled to run after a specific time while retaining the credit balance of the VM execution time slice such that resources are actually being stolen. Modified versions of scheduler (Zhou et al., 2011[80]) may improve security of hypervisors while maintaining fairness and efficiency. Some kinds of attacks to the VM scheduler are covert channel attacks, priority manipulation attacks and resources starvation attacks.

---

[75] Riquet, D., Grimaud, G., Hauspie, M., 2012. "*Large-scale coordinated attacks: impact on the cloud security*". In: 2012 Sixth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), pp. 558–563. http://dx.doi.org/10.1109/IMIS.2012.76.

[76] Scarfone, K.A., Mell, P.M., 2007. Sp 800-94. "*Guide to Intrusion Detection and Prevention Systems (IDPS)*". Technical Report, Gaithersburg, MD, United States

[77] A *botnet* is a group of Internet-connected devices, each of which runs one or more bots. Botnets can be used to perform distributed denial-of-service (DDoS) attacks, steal data, send spam, and allow the attacker to access the device and its connection. The owner can control the botnet using command and control (C&C) software. The word "*botnet*" is a portmanteau of the words "robot" and "network". The term is usually used with a negative or malicious connotation.

[78] McMillan, R., 2009. "*Hackers Find a Home in Amazon EC2's Cloud*". URL : ⟨http://www.computerworld.com/article/2521744/security0/hackers-find-a-home-in-ama zon-s-ec2-cloud.html⟩.

[79] Two techniques can be used to detect sniffer programs. First, the *Address Resolution Protocol* (ARP) detect sniffer by sending trap ARP packets (contains fake hardware addresses) to a suspicious host. Second, *Round-Trip Time* (RTT) make use of samples of RTT measurements of ICMP packets and then by using a statistical model make a probabilistic decision.

[80] Zhou, F., Goel, M., Desnoyers, P., Sundaram, R., 2011. "*Scheduler vulnerabilities and coordinated attacks in cloud computing*". In: 2011 10th IEEE International Symposium on Network Computing and Applications (NCA), pp. 123–130.

A *covert channel* is a communication channel that allows entities to transfer information in a way that violates the system's security policy. While these channels can increase the privacy and security of critical communication, criminals use covert channels to carry out cyberattacks. To achieve power-efficient computing, processors engage idle power management mechanisms to turn on/off idle components according to the dynamics of the workload. A processor’s hardware components are classified and managed through the core and the *uncore*. The uncore is the supporting hardware shared by the cores, hence the decision to turn it on/off depends on the cores’ activities. Such dependency implies a covert channel threat in multi-core platforms. Specifically, the power status of the uncore reflects the workload pattern of the active core, and it can be probed by any process running on the processor. This allows the process to infer the workload information of the active core. As such, this covert channel can work across processors and violate VM isolation.

In *priority manipulation attacks*, attackers may manipulate the input and output buffers of victim tasks, cause instability, and/or steal hidden information. The success of those attacks highly depends on the proper positioning of the malicious task with respect to the victim tasks in the actual schedule, changing the priority ranking of the tasks in the VM scheduler. By overloading the system with high-priority tasks or exploiting scheduling algorithms, they can starve legitimate tasks of resources, causing delays or disruptions. This manipulation allows attackers to gain unauthorized access, extract sensitive data, or degrade the performance of co-resident VMs, compromising the cloud environment's integrity and availability. Then, the malicious task executes the necessary actions to steal hidden information.

The *resources starvation (RS) attack* is a type of DoS with the main purpose to consume as much as possible resources on the target system(s) with relative light load on the attacking system. To make this happen, Resource Starvation (RS)/Denial of Service (DoS) attack at the VM level exploits shared resources like CPU, memory, or disk I/O. Attackers overload the VM by consuming excessive resources, starving legitimate tasks and causing performance degradation or service unavailability. This disrupts the VM's functionality and can impact co-resident VMs, compromising the cloud environment's stability and availability.

4.2.6.14 Data scavenging

While erasing data from a storage device, the file systems do not remove data completely. Consequently, the removed data may be recovered by attackers which is referred to as data scavenging (Hashizume et al., 2013; Jansen, 2011; Ertaul et al., 2010; Grobauer et al., 2011; Sen, 2013; Townsend, 2009). Various techniques to counter data scavenging are reported in AWS (2014). This attack is also considered as multilayer attack (affects secondary PaaS and SaaS) but can be also categorized in the *Storage layer*[81] in Cloud Computing.

4.2.6.15 Data deduplication

[81] The Storage Layer in cloud computing is responsible for storing, managing, and retrieving data across distributed cloud environments. It provides scalability, reliability, and accessibility, ensuring that data is securely stored and efficiently accessed by applications and users.

With data deduplication being performed for minimizing storage and bandwidth requirements, it becomes possible to identify the location of the files and their contents (Harnik et al., 2010). The term “data deduplication” refers to techniques that store only a single copy of data and provide links to that copy instead of storing other actual copies of this data. With the transition of services from tape to disk, data deduplication has become a key component in the backup process.

The attacker carefully crafts data that collides with original data of a legitimate user (in terms of hashing). This enables the attacker to obtain a link to the original data and thus access them. In another example, the attacker uploads a huge collection of data that bypasses deduplication in order to cause excessive resource consumption. We are refenced here, in the IaaS layer, since we talk about the use of data resources.

By storing and transmitting only a single copy of data, deduplication offers savings of both disk space and network bandwidth. However, there is an inherent risk in entrusting data in the storage cloud, since the data owner is basically releasing control over his/her data. Yet, a wide range of users and applications are more than willing to hand over their data storage tasks to a cloud provider. They put their trust in the integrity of the cloud provider and in the security of the access control mechanisms that it uses. The deduplication exploitation risk may be mitigated by ensuring the deduplication to occur only if there is a specific number of copies of the file (Kaaniche and Laurent, 2014). This attack is also considered as multilayer attack (affects secondary the PaaS and the SaaS) but can be also categorized in the *Storage layer* in Cloud Computing.

4.2.6.16 Passive eavesdropper

The eavesdropper can wiretap the repair process of a subset of nodes in the storage system (actually in DSS-Distributed Storage System). Distributed storage is the default technique for storing data in all new generation applications. The data from a file is stored in a decentralized manner on several unreliable nodes/disks that when collectively used can recover the entire file. To ensure robustness to disk failures, the simplest scheme is to replicate and store the data across several disks. Securing sensitive data, such as personal and financial records, from eavesdroppers is necessary to ensure data privacy for the users. Hence, a DSS should be secure apart from satisfying the reconstruction and regenerating requirements. This attack is also considered as multilayer attack and can be also categorized in the *Network layer* in Cloud Computing.

4.2.6.17 Steganography attacks

With a steganography attack, the attackers embed malicious code within files being transmitted over the network. Schemes like StegAD[82] may be used to identify the steganographied files and possible locations of injected code in them. Steganography attacks can include or not the injection of malware. As they can be used for other purposes like data exfiltration, data and control communication, watermark removal, etc. Steganography enables the attackers to evade some

[82] *StegAD* was proposed, which is a novel scheme for defending Audio steganography Attacks. StegAD consists of the enhanced-RS algorithm and the SADI algorithm. The enhanced-RS algorithm is first applied to detect the audio steganographied files, and after that, SADI is applied to interfere in possible hiding place with random noise and later compensate the damage. To evaluate the performance of StegAD, a prototype cloud storage system was built based on which extensive experiments were conducted in terms of detecting steganographied files, interference effects on various steganography techniques and audio quality loss.

security tools, including network traffic scanners. This attack is also considered as multilayer attack (affects secondary the PaaS and the SaaS) and can be also categorized as an attack in the *Network layer, Application layer, Storage layer* and in *Virtualization layer*[83] in Cloud Computing.

4.2.6.18 VM migration attacks

When an active VM is migrated from the host physical machine to another physical machine, the contents of the VM files become vulnerable to various attacks. Since one of the biggest benefits of IaaS to the customer is the rapid elasticity of their provision. This elasticity creates new VMs on one physical machine. Whilst such creation is transparent and potentially seamless, it may also introduce vulnerabilities. *Virtual Machine (VM) Elasticity* refers to the cloud’s ability to dynamically scale computing resources up or down by creating or terminating VMs based on demand. While this improves resource utilization and cost efficiency, it also introduces new security vulnerabilities[84] that attackers can exploit. Under this attack, a malicious insider may exploit vulnerabilities associated with the creation of new VMs to obtain large volumes of cloud-user data. Alternatively, the log of execution state being maintained for implementing a rollback may become accessible during migration. These attacks can be considered that affects, also, the *Network layer*[85] in Cloud Computing, secondary.

4.2.7 Cases confirming the necessity of protecting the cloud privacy and security. Consequences and thoughts

4.2.7.1 Security attacks on SaaS cloud layer

A recently published report found that 68% of enterprises consider SaaS misconfigurations as the biggest threat to their cloud security (InfoSecurity Magazine, 20 May 2021). Another research from Cloud Security Alliance, 2021 State of Cloud Security, indicates that security misconfigurations are the main contributor for 22% of security incidents. Unfortunately, these SaaS misconfigurations can lead to severe repercussions.

Under application vulnerabilities, developers who build applications using SaaS components may not use secure coding practices, resulting in security vulnerabilities within applications. This can affect not only the security of the individual application but also the platform that hosts it and potentially other platforms that are co-hosted with it.

[83] The Virtualization layer in cloud computing is responsible for abstracting physical hardware and enabling multiple virtual environments to run on a single physical machine. It acts as an interface between the physical infrastructure (servers, storage, and networks) and the virtual instances (VMs, containers, or cloud services) that operate on top of it.

[84] As the "VM sprawl", where many instances run without proper tracking or security enforcement, or “Insecure Auto-Scaling Configurations” creating new VMs without proper security controls.

[85] In cloud computing, the Network Layer refers to the part of the cloud architecture responsible for managing the network infrastructure, which facilitates communication between different components of the cloud environment (such as virtual machines, storage, applications, and users). The network layer handles the routing, switching, and management of traffic across different resources, ensuring that data can be sent and received securely and efficiently.

The list of possible misconfigurations in SaaS, whether intentional or by mistake, can be endless. Further, there are some exploited misconfigurations that are being exploited repeatedly over time. Here is an example from a real-world attack[86].

1. U.S. Treasury Department

In May 2009, the U.S. Treasury Department discovered that malicious code was present on their website and reportedly shut down four public-facing sites of the Bureau of Engraving and Printing. Subsequent investigations identified the cause due to a third party. Hackers had added a small snippet of virtually undetectable *iframe HTML code* that redirected visitors to a web site in the Ukraine that then launched a variety of Web-based attacks based on a commercially available attack-kit called the Eleonore Exploit pack. Similarly, the chief research officer of Anti-Virus Guard (AVG) technologies discovered that the affected pages are vulnerable to malicious code.

4.2.7.2 Security attacks on PaaS cloud layer

PaaS platforms are attractive targets for hackers because they provide access to a wide range of applications and data. As such, hackers can exploit various vulnerabilities to fulfil their purposes, like the following:

PaaS misconfigurations

Platform vulnerabilities—attackers can exploit vulnerabilities in the platform to gain control over the system or steal sensitive data.

Limited visibility—as the complexity of the PaaS infrastructure increases (ie complexity of compliance or complexity of building and maintaining the underlying infrastructure), it becomes more difficult to achieve visibility, detect vulnerabilities and threats and understand how to mitigate them.

Here are examples from real-world attacks on PaaS cloud layer[87]:

1. Equifax attack

In 2017, the credit reporting agency *‘Equifax’,* one of the largest human intel databases in the world, was the victim of a *man-in-middle* (MiTM) attack due to an unpatched vulnerability in its web application framework called Apache Struts that was running on a server in their DMZ (demilitarized zone), facing the Internet. What we do know is that the Internet-facing applications are at the extreme risk from external attackers, and that is why they are placed into the so-called demilitarized zone (DMZ), a low-trust segment of the network separated from the internal company network. Unfortunately, the attack was sophisticated, so it spread over other parts of the Equifax enterprise system (initial user credential information exposed was utilized to attack on other system parts). As such, it managed to expose the financial information of nearly 150 million people.

What makes this significant is that during the congressional hearing that followed this breach, it was revealed that penetration testers had found this vulnerability months before. As for their official report, the breach happened in mid-May of 2017. There was a patch available for this

[86] www.arcserve.com, 7 Most Infamous Cloud privacy breaches, 20 Dec 2023

[87] https://proton.me/blog/kazakhstan-internet-surveillance, https://archive.epic.org, www.aquasec.com

vulnerability in March. That means two months passed after they knew they were vulnerable and did nothing to fix it.

2. Government of Kazakhstan

In 2015, the government of Kazakhstan created a root certificate which could have enabled a man-in-the-middle attack on HTTPS traffic from Internet users in Kazakhstan. The government described it as a "national security certificate". If installed on users' devices, the certificate would have allowed the Kazakh government to intercept, decrypt, and re-encrypt any traffic passing through systems it controlled.

In July 2019, Kazakh ISPs started messaging their users that the certificate, now called the Qaznet Trust Certificate, issued by the state certificate authority the Qaznet Trust Network, would now have to be installed by all users. Sites operated by Google, Facebook and Twitter appeared to be among the Kazakh government's initial targets.

On August 21, 2019, Mozilla and Google simultaneously announced that their Firefox and Chrome web browsers would not accept the government-issued certificate, even if installed manually by users. Apple also announced that they would make similar changes to their Safari browser. As of August 2019, Microsoft has so far not made any changes to its browsers, but reiterated that the government-issued certificate was not in the trusted root store of any of its browsers, and would not have any effect unless a user manually installed it.

In December 2020, the Kazakh government attempted to re-introduce the government-issued root certificate for a third time. In response to this, browser vendors again announced that they would block any such attempt by invalidating the certificate in their browsers.

3. Kaseya

In July 2021, IT solution provider Kaseya experienced an attack on its remote monitoring and network perimeter security tools. It was a supply-chain based ransomware attack, designed to gain administrative control over Kaseya services and use them to infect the networks of managed service providers and their customers. The attack took down the company’s SaaS servers and affected on-premises virtual SAN appliances (VSA) used by Kaseya customers in 10 countries. Kaseya was proactive in responding to the attack and alerted customers immediately. Later, the company deployed a VSA detection tool to allow its customers to analyze VSA services and identify signs of vulnerabilities.

4. Facebook

Facebook was breached sometime before August 2019 but decided not to notify over 530 million users that their personal data was stolen until April 2021. An app was designed to collect data from Facebook and users voluntarily participated in the quiz. However, due to misconfigured access controls in Facebook's specific app, it was able to collect not only the data of the users who took the quiz but also the data of their friends—without their explicit consent. The data included phone numbers, full names, locations, some email addresses, and other details from user profiles.

While Facebook later posted an account about the attack on its blog, the damage to the company's reputation was tainted. Facebook says it found and fixed the issue immediately, but the ripple effect even hit founder Mark Zuckerberg. He had to answer to federal regulators to settle a privacy case with the Federal Trade Commission that included a $5 billion penalty paid by the

company. Things only worsened in October 2021 when whistleblower Frances Haugen claimed that Facebook chooses profits over safety.

5. Alibaba

While the hacker did not get ahold of encrypted information like passwords, the breach was severe enough that the company notified the police. As it happened in China, the full consequences of this attack will likely never be made public. However, it is an example that makes a strong case for better monitoring systems and networks. The issue is that potentially the information exfiltration was done very carefully in order not to cause suspicions.

In a nutshell, the vulnerabilities (a privilege escalation flaw in AnalyticDB and a remote code execution bug in ApsaraDB RDS) made it possible to elevate privileges to root within the container, escape to the underlying Kubernetes node, and ultimately obtain unauthorized access to the API server. Armed with this capability, an attacker could retrieve credentials associated with the container registry from the API server and push a malicious image to gain control of customer databases belonging to other tenants on the shared node.

6. Sian Weibo

Sina Weibo is one of China's largest social media platforms. In June 2020, the personal details of more than 538 million users—including real names, site usernames, gender, and location—and phone numbers for 172 million users were posted on the dark web and other places. While it is not clear how the incident originated, the hacker put Weibo's data up for sale for a mere $250, most likely because it did not include passwords.

The attackers targeted specific weaknesses in Sina Weibo’s data management and security systems. These vulnerabilities allowed unauthorized access to user data that was supposed to be protected. The breach was not a direct ‘hack’ in the traditional sense but rather a sophisticated exploitation of system loopholes. Even though Weibo is heavily monitored and censored these days, it is still used, at times, to share unfiltered news from around the country. As a result, anonymous Weibo users may face the most significant risks due to the possibility of new breaches. The method used was *data scraping*, where automated scripts are employed to collect publicly available information at a scale not feasible for individual users. This approach, while not hacking in the conventional sense, bypasses normal user privacy settings and protections, raising questions about the adequacy of existing security measures against such methods.

7. Code Spaces

A real-world example of a cloud attack related to Platform-as-a-Service (PaaS) is the Code Spaces breach in 2014. Code Spaces was a software development platform that provided version control and project management tools, hosted on Amazon Web Services (AWS). The attack led to the company's shutdown and serves as a cautionary tale for PaaS security.

How the Code Spaces Breach Happened:

Initially, attackers gained access to Code Spaces' AWS management console through compromised credentials, likely obtained via phishing or brute-force attacks. Then, the attackers demanded a ransom in exchange for not destroying the company's data. When Code Spaces refused to pay, the attackers escalated their actions. Using the compromised credentials, the attackers deleted critical data and backups stored in Code Spaces' AWS Elastic Block Store (EBS) and Simple Storage

Service (S3) buckets. They also deleted snapshots, virtual machine instances, and other infrastructure components, effectively crippling the platform. Code Spaces lost the majority of its data and infrastructure, making it impossible to recover operations. The company was forced to shut down permanently.

This is a PaaS-related attack, because it has a PaaS dependency, as Code Spaces relied on AWS as its underlying PaaS provider for hosting its development tools and data storage. The breach exploited weaknesses in how Code Spaces managed its PaaS environment. Moreover, there were misconfigured access controls, because the attackers exploited weak access management, such as insufficient multi-factor authentication (MFA) and overly permissive credentials, to gain control of the AWS console. Finally, there was a lack of backup security, because the company's backups were stored in the same PaaS environment without proper isolation or redundancy, making them vulnerable to deletion.

Some lessons learned: *Strengthen Access Controls* as implemented MFA and enforce least-privilege access for PaaS management consoles, *Isolate Backups* by storing backups in a separate, secure environment to prevent them from being compromised during an attack and *Monitor for Unusual Activity* by using logging and monitoring tools to detect unauthorized access or changes to PaaS resources.

The Code Spaces breach highlights the risks of inadequate security practices in PaaS environments and the devastating consequences of failing to protect cloud infrastructure.

8. Capital One

Another real-world example of a cloud attack related to Platform-as-a-Service (PaaS) is the 2019 Capital One data breach. This incident involved a misconfigured web application firewall (WAF) in a PaaS environment hosted on Amazon Web Services (AWS). The breach exposed sensitive data of over 100 million customers.

How the Capital One Breach Happened:

Initially, an exploitation of misconfigured WAF took place. A former AWS employee, Paige Thompson, exploited a misconfigured web application firewall (WAF) in Capital One's AWS environment. The WAF was part of a PaaS setup used to host the company's applications. The misconfiguration allowed Thompson to access a Server-Side Request Forgery (SSRF) vulnerability, which enabled her to send unauthorized requests to the AWS metadata service. Then, using the SSRF vulnerability, Thompson gained access to AWS Identity and Access Management (IAM) credentials stored in the metadata service. These credentials allowed her to access AWS S3 buckets containing sensitive customer data, including names, addresses, credit scores, and Social Security numbers. Thompson exfiltrated over 100 GB of data from Capital One's S3 buckets and later shared some of the stolen information online. The breach was detected when Thompson boasted about her actions on GitHub and Slack. She was subsequently arrested and charged with computer fraud and abuse.

This is a PaaS-related attack, due to : (a) PaaS Misconfiguration: the breach stemmed from a misconfigured WAF in Capital One's PaaS environment, highlighting the risks of improper configuration in cloud platforms, (b) SSRF Vulnerability: the attack exploited a common vulnerability in cloud environments, where metadata services can be accessed if not properly secured, and (c) Shared Responsibility Model: the incident underscored the importance of the shared

responsibility model in cloud security. While AWS provided the infrastructure, Capital One was responsible for securing its applications and configurations.

Some Lessons Learned: *Secure Metadata Services* by ensuring metadata services which are not exposed to unauthorized access, *Regular Configuration Audits*, by conducting regular audits of PaaS configurations to identify and fix vulnerabilities, and *Monitor for Unusual Activity* by implementing robust monitoring and logging to detect unauthorized access or data exfiltration attempts.

The Capital One breach demonstrates how misconfigurations in PaaS environments can lead to significant data breaches, emphasizing the need for proper security practices in cloud platforms.

4.2.7.3 Security attacks on IaaS cloud layer

A few examples of attacks specific to IaaS services are denial of service (DoS) attacks against cloud-based computer resources and hijacked IaaS resources used to mine cryptocurrencies FirEye case below). DoS attacks can last from a few hours to a much more time, costing companies and consumers time and money while their resources and services are unavailable. Here are examples from real-world attacks on the IaaS cloud layer.

1. AWS

One of the recent examples of a DDoS Attack occurred in February 2020, impacting Amazon Web Services (AWS), a cloud computing service used by over a million companies, individuals, and government entities. The hackers used directories on Connection-less Lightweight Directory Access Protocol (CLDAP) servers to send huge amounts of information to AWS's servers—as many as 2.3 terabits per second (Tbps). However, Amazon was able to stop the attack before it became a big security risk for its customers (by *AWS Shield* which is a managed threat protection service that safeguards applications running on AWS against exploitation of application vulnerabilities, bad bots, and Distributed Denial of Service (DDoS) attacks).

2. Dyn

In October 2016, a well-targeted DDoS attack was carried out on a domain name system (DNS) provider, Dyn, which hosts and manages the domain names of select companies on its servers. Due to that attack, some of the Dyn’s servers in specific regions (North America & Europe) were compromised and this also affected the websites of the companies these servers hosted. In particular, the attack on Dyn flooded its servers with overwhelming traffic, creating a massive web outage and shutting down over 80 websites, including major sites like Twitter (now X), Amazon, Spotify, Airbnb, PayPal, and Netflix. The Dyn attack is related to the cloud because it exploited cloud-connected devices and targeted a critical cloud infrastructure component (DNS), disrupting numerous cloud-based services and highlighting the need for improved security in IaaS and IoT ecosystems.

3. FireEye

FireEye disclosed in December that they had become the victims of a Nation-State attack. They described the breach as different from any of the tens of thousands of attacks they have responded to over the past 25 years, showing that new attack methods, or combinations of attack methods, are appearing all the time. FireEye’s red teaming and penetration testing tools were targeted and stolen in the attack, the same ones that the company uses to test their customer

networks. These tools have yet to be leaked or utilized elsewhere but could be used to uncover vulnerabilities against future targets or simply exposed to discredit FireEye directly.

The cybersecurity company FireEye announced (in 2020) in a blog post and SEC filings that they had been hacked. FireEye is one of the world’s largest cybersecurity firms with over 9,600 customers across 103 countries, including more than 50% of the Forbes Global 2000. As reported by the company, the hack was highly sophisticated and likely state sponsored. Initial reports suspect the Russian SVR intelligence service APT 29 (Advanced Persistent Threat Group 29) but the investigation continues. The FBI and Microsoft Corp have been called in to help with the investigation.

The attackers gained access to FireEye’s Red Team toolkit which is a set of scripts, tools, scanners, and techniques used by FireEye to test client security postures. The breach resulted in a set of proprietary “Red Team” tools used with FireEye’s Commando Virtual Machine (VM) being stolen. Commando VM is a penetration testing tool purpose-built for exploiting Windows devices. The tool leverages a list of known vulnerabilities to gain network access and identify flaws in FireEye customers’ security. Some of the toolkit components were already released in the public domain and widely available but some were still proprietary in-house tools used by the Red Team. The Red Team tools stolen did not contain any zero-day exploits. The tools apply well-known and documented methods that are used by other red teams around the world. It is important to note that FireEye has not seen these tools disseminated or used by any adversaries, and they will continue to monitor for any such activity along with their security partners. In response to the attack, FireEye has released countermeasure tools to the community to help identify attacks using the Red Team toolkit and defend against attacks.

Conclusively, the attackers gained access to FireEye's network by exploiting compromised credentials (likely through phishing or password spraying) of employees with access to critical systems. They used these credentials to bypass multi-factor authentication (MFA) by exploiting vulnerabilities in the authentication process, possibly through token theft or session hijacking. Then, attackers moved laterally across FireEye's network using legitimate tools and scripts (a technique known as "living off the land") to avoid detection. They escalated privileges to gain access to highly sensitive systems, including those storing Red Team tools. Attackers exfiltrated FireEye's proprietary Red Team tools, which could be repurposed for malicious activities. However, FireEye stated that the stolen tools did not contain zero-day exploits. They also accessed and stole sensitive data related to FireEye's government and private sector clients. Finally, the attackers used advanced anti-forensic techniques to cover their tracks, making it difficult for FireEye to detect the breach initially and operated with a high level of sophistication, mimicking legitimate user behavior to evade detection by FireEye's own security tools.

5 *Cloud privacy protection approaches and countermeasures*

In this chapter, we present the countermeasures regarding cloud privacy and specifically the attacks that they address. Then, we supplied well-designed criteria for classifying the analyzed countermeasures for cloud privacy attacks. After that, we propose for every category the countermeasures which satisfy the majority of the selected criteria that could be applied for the protection of cloud privacy from organizations and individuals, in different cloud environments. Finally, we propose a *countermeasures selection algorithm* for a company to choose the proper countermeasures to protect cloud privacy from attacks.

5.1 Analysis and taxonomy of the approaches and countermeasures for attacks and threats, regarding cloud privacy and security

In the following sections (5.1.1 – 5.1.5), we analyze and present the main countermeasure’s categories for attacks and threats regarding the protection of cloud privacy and security, analyze every countermeasure inside each category for their best use, their strengths and weaknesses and supply an overall evaluation value. The categories are: MiTM, IDS, DLP, Data storage and Data privacy protection.

5.1.1 Man-in-the-Middle (MiTM) attacks: countermeasures for cloud privacy and security

5.1.1.1 MiTM Symptoms and detection

Below is an analysis of symptoms, detection techniques, and countermeasures to mitigate these threats. Identifying a MiTM attack requires recognizing signs of intercepted or manipulated communication. Common symptoms, include:

1. Network Anomalies
    - Indicators: Unusual latency, sudden slowdowns, or unexpected network behavior.
    - Example: Unexpected routing of traffic to unknown IP addresses or geolocations.
2. Certificate Warnings

- Indicators: Alerts for invalid, untrusted, or mismatched SSL/TLS certificates.
- Example: A browser warning like *“This site’s certificate cannot be verified.”* This suggests the use of forged certificates by an attacker.

3. Unusual Application Behavior
   - Indicators: Disrupted cloud application performance, such as repeated logouts or authentication errors.
   - Example: Inconsistencies in transmitted data, such as corrupted or missing files during cloud uploads.
4. Unauthorized Traffic Redirection
   - Indicators: Unexpected redirections to phishing[88] websites or malicious domains.
   - Example: DNS spoofing causes users to visit attacker-controlled websites resembling legitimate services.
5. Repeated Authentication Issues
   - Indicators: Unexplained login failures or account lockouts.
   - Example: Attackers attempting to reuse intercepted credentials disrupt legitimate user sessions.

Detecting MiTM attacks involves monitoring network traffic, validating certificates, and analyzing user behavior. Key detection techniques include:

1. Packet Analysis
   - Tools like Wireshark can analyze network traffic for irregularities, such as duplicate packets, unusual headers, or encrypted traffic anomalies.
2. Certificate Validation
   - Enforcing TLS certificate pinning ensures communication only with predefined, trusted certificates. Untrusted certificates trigger immediate alerts.
3. Intrusion Detection and Prevention Systems (IDS/IPS)
   - Signature-based detection: Identifies known attack patterns.
   - Anomaly-based detection: Monitors traffic for deviations from established baselines.
4. DNS Monitoring
   - Detects unauthorized DNS changes or spoofing attempts using DNSSEC (Domain Name System Security Extensions).
5. Behavioral Analytics
   - AI-driven solutions analyze user behavior to detect anomalies, such as access from unusual locations or sudden spikes in data requests.

5.1.1.2 MiTM countermeasures

Organizations can employ various countermeasures to mitigate MiTM risks effectively:

[88] The “Intelligent Cloud Based Email Encryption and Decryption System” (ICLEEDS) (Ayodele and Adeegbe, 2013) is a framework to improve the security of cloud-based email messages. This intelligent machine learning encryption system helps the systems to improve and protect users email from any attack such as phishing, spoofing, relaying of previous message etc.

1. Enforce End-to-End Encryption
   - Encrypt sensitive data at rest within the cloud environment.
2. Implement Strong Authentication Mechanisms
   - Multi-Factor Authentication (MFA): Adds an additional layer of protection by requiring multiple forms of (user identity) verification.
   - Zero Trust Architecture: Restricts access based on continuous identity verification and device posture.
3. DNS Security Enhancements
   - Use DNSSEC to validate the authenticity of DNS records.
   - Employ secure DNS resolvers to mitigate risks of redirection attacks.
4. Secure Network Configurations
   - Implement VPNs (Virtual Private Networks) for encrypted remote access to cloud services.
   - Use firewalls and network segmentation to limit exposure to attack.
5. Use of AI-based/ML-based IDS/IPS
   - Deploy AI solutions to identify unusual communication patterns and proactively mitigate threats.
   - Behavioral analytics can flag suspicious user activities or connection anomalies.
6. Employee Training
   - Educate employees on recognizing phishing attempts, certificate warnings, and the importance of verifying secure connections.
   - Regularly conduct security awareness sessions.

MiTM attacks represent a significant threat to cloud privacy, particularly in environments relying on extensive data exchange and remote accessibility. Proactive detection techniques and robust countermeasures, such as end-to-end encryption, certificate validation, and AI-driven monitoring, are essential to mitigate these risks. By integrating these practices into a comprehensive security strategy, organizations can safeguard cloud environments against privacy threats.

Regarding the ‘*XML signature wrapping’* (XSW) attack, a system that can tolerate this attack can rely on the use of the UNWRAP method (Nasridinov et al., 2012[89]). The authors first construct the SOAP message header in a specific manner (building SOAP message elements structure using ontology). Since they construct the structure of the SOAP message element, then they attached to the SOAP message header. If the attacker modifies the SOAP message, it is detected by observing the change in the relationship of the signed elements. SOAP messages are also written as a log by using this method. If there is a security breach later, the log can be examined so as to find its root cause.

Below, in Table 5 we present our proposed countermeasures for the MiTM category, with their best use, strengths and weaknesses. The proposed countermeasures for the MiTM category, have been resulted from our research on the literature (see also Ch. 3, Methodology review), which we consider them as the most used.

[89] Nasridinov, A., Byun, J.-Y., Park, Y.-H., 2012. “*UNWRAP: An approach on wrapping- attack tolerant SOAP messages*”. In: Proceedings of the Cloud and Green Computing (CGC), 2012 Second International Conference on. 2012. IEEE.

Table 5: MiTM countermeasures category & their best use cases

Countermeasure	Best For	Strengths	Weaknesses	Value
End-to-End Encryption (HTTPS, VPNs, IPsec, SSH)	Preventing eavesdropping on web, email, and network traffic	Ensures data confidentiality	Metadata Leakage – While E2EE protects content, metadata (e.g., IP addresses, timestamps, sender/receiver details) remains exposed, which can be analyzed for patterns.	*High* (Encrypts data before transmission, ensuring only intended recipients can decrypt)
Certificate Pinning	Securing mobile apps and preventing fake certificates	Prevents MITM via rogue CAs	Requires updating if the certificate changes	*High* (Prevents attackers from using fake certificates by validating expected certificates)
Mutual Authentication (Client & Server Authentication, MFA)	Ensuring only trusted devices and users connect	Defends against impersonation	Implementation complexity	*High* (Adds an extra layer of authentication, making unauthorized access harder)
DNS Security (DNSSEC, DoH, DoT, Encrypted DNS)	Protecting against DNS spoofing and hijacking	Prevents attackers from redirecting traffic	Not all DNS resolvers support it	*Moderate* (Protects against DNS spoofing but not direct MiTM attacks)
Network Security (Firewall, NAC, ARP Inspection, VLAN Segmentation)	Protecting local networks from ARP spoofing and packet interception	Blocks unauthorized traffic	Configuration complexity	*Moderate* (Monitors and blocks suspicious activity, but doesn't fully prevent MiTM)
Zero Trust Architecture (ZTNA, Identity-Based Access)	Preventing lateral movement in networks	Reduces attack surface	Requires strong identity & access management	*High* (Assumes no trust by default, verifying identity and device security before access)
AI-Powered Threat Detection	Identifying abnormal behavior & anomalies	Detects new and evolving attack patterns	May generate false positives	*High* (Uses AI/ML to detect suspicious patterns and potential MiTM attacks in real time)
Public Key Infrastructure (PKI) & Digital Signatures	Authenticating email, documents, and messages	Prevents tampering and impersonation	Key management complexity	*High* (Ensures strong authentication and integrity using

				cryptographic certificates)
Basic SSL/TLS Encryption (No Pinning, No Certificate Validation)	Non-sensitive communications, low-risk environments	Easy to implement - Provides basic encryption and confidentiality during transmission	Vulnerable to MiTM attacks if certificates are compromised	*Low* (without certificate pinning or proper certificate validation, this method is vulnerable to MiTM attacks)
UNWRAP method	XML signature wrapping attack	Bypasses Signature Verification Without Breaking It, No breaking encryption	Can Be Prevented with Strong Reference Validation, Requires XML Interception	*High* (Aims to provide a more secure authentication framework, reducing MiTM vulnerabilities)

5.1.2 Intrusion detection system (IDS) in cloud computing

5.1.2.1 Intrusion detection in cloud security and privacy

We have thoroughly explored that each cloud layer suffers from certain vulnerabilities. These vulnerabilities are introduced by different programming and configuration errors of Internet and web applications, wireless communication networks of users or the CSPs. There are also issues concerning improper resource isolation, the wrong design of protocols or applications and issues related to user mistakes (e.g., weak passwords).

Numerous services are provided by the CC to the end users. Therefore, to provide safe and reliable services, assurance is quite important. The IDS is one of the practical approaches to defend against these attacks and address various types of vulnerabilities. In the cloud multi-tenant environment, IDS provides the best way to defend against known and unknown attacks.

The attacker can attack by taking advantage of vulnerabilities in VM environment and can sabotage the physical host and the virtual networks supported by the host. However, it gets even worse when an attacker launches the attacks on multiple virtual hosts over one single physical machine. If the attacker compromises the hypervisor or the physical host, this could be easily done. So, it is the outcome of the attack of the attacker - compromising a hypervisor leads to compromising the VMs managed by the hypervisor. The IDS can help to reduce the risk of this type of attack. For the cloud environment hypervisor-based security solutions are effective to protect the virtualized environment.

Thus, the Intrusion Detection System (IDS) is an essential component of a defensive mechanism to protect network and computer systems (as those involved in cloud environments) against various attacks at different (cloud) service models. It can detect (D)DoS attacks as well as abnormal behaviors. However, if, for instance, a user accidentally opens a backdoor and then intruders invade the system and steal cleverly sensitive information, the IDS cannot do anything

about that. In the context of sensitive data, DLP software might be better as it can detect and prevent threats related to data privacy (see also Section 4.2.2).

An IDS involves the procedure of monitoring network traffic (among others) to detect any vulnerability exploits against a target application or system. Thus, the main task of intrusion detection is to enhance security in a computer network. Basically, an IDS is a whole compound that includes a component dedicated to intrusion detection. Even more, IDS may contain a variety of flavors and approaches to detect suspicious traffic in different ways (in this case, it is a hybrid IDS comprising multiple detection components).

When an IDS detects an intrusion or suspicious activity, it generates an alert or notification for security administrators. These alerts provide information about the nature of the incident, the affected system or network, and any additional details to aid in the response and mitigation process

There is continuous evolvement in the current enterprise architecture. An intrusion detection system is the procedure which monitors network traffic to detect any vulnerability exploits against target applications or systems. Thus, the main task of intrusion detection is to enhance the security in a networked computer environment. Basically, the IDS is a whole compound that includes various components. This IDS system contains a variety of flavors, and approaches detect suspicious traffic in different ways.

It is important to note that IDS is not a standalone solution but works in conjunction with other security measures like firewalls, antivirus software, and security policies. Additionally, an extension of IDS called intrusion prevention systems (IPS) ais often used to not only detect but also actively block or prevent detected threats.

As every cloud service delivery model is different from others service models similarly, IDS techniques used for each cloud services model also differ. The Intrusion Detection System (IDS) is an essential component of a defensive mechanism to protect network and computer systems against various attacks at different service models.

Traditional firewalls methods cannot provide a good solution as the firewall rules need to be modified for each trivial update in enterprises. A comparison between IDS and firewall is shown below (Table 6). Moreover, the approaches so far proposed for IDS in cloud environment are also discussed (Fig. 14).

Table 6. Differences between IDS and Firewall (Octay, 2013)

Differences between IDS and Firewall		
Criteria	IDS	Firewall
Working Principle	It monitors real-time traffic for attack patterns, generating alerts when suspicious activity is detected	It filters traffic based on IP addresses and port numbers, acting as a barrier between trusted and untrusted networks using predefined security rules
Function	It identifies and alerts administrators of potential threats without taking direct action	Monitors and manages traffic according to pre-established security rules, preventing unauthorized access
Location	It is usually located within the internal network to monitor traffic without interfering with data flow.	It is typically placed at the network perimeter, serving as the first line of defense against external threats.
Traffic Filtering	It examines traffic behavior to detect malicious patterns and generates alerts	It filters traffic based on pre-established rules without analyzing traffic patterns
Performance Impact	Can significantly impact performance depending on the complexity of its traffic analysis	Minimal impact on network performance since it uses simple rules to filter traffic

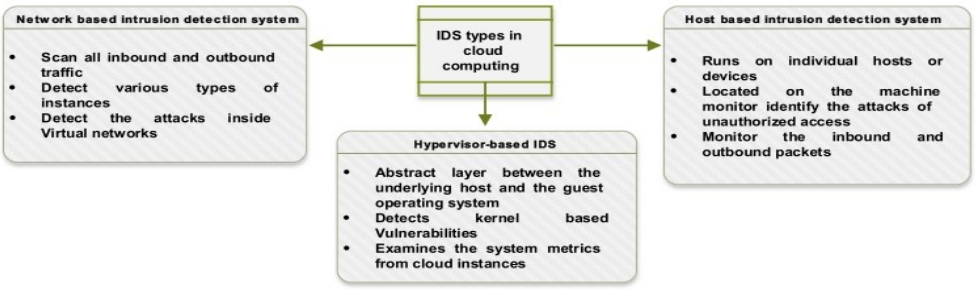

Fig 14. Types of IDS in CC.

5.1.2.2 Intrusion detection approaches

We analyze the approaches proposed for IDS along with their pros and cons (Oktay and Sahingoz, 2013), as follows:

- Signature-based detection
- Anomaly-based detection
- Artificial neural network (ANN) based IDS
- Fuzzy logic-based IDS

Signature-based detection. Signature-based detection, also known as rule-based detection, relies on pre-defined signatures or patterns of known attacks to identify intrusions. Signature intrusion detection systems (SIDS) are based on pattern matching techniques to find a known attack; these are also known as Knowledge-based Detection or Misuse Detection (Khraisat et al., 2018).

In SIDS, matching methods are used to find a match between the current packet or behaviour and a known intrusion. In other words, when the packet/behaviour signature matches that of an existing/known intrusion, an alarm signal is triggered. For SIDS, the host's logs are inspected to find sequences of commands or actions which have previously been identified as a malware intrusion. SIDS system usually attains excellent precision detection for known intrusions (Kreibich & Crowcroft, 2004).

Anomaly-based detection. The anomaly detection technique looks up into the baseline (normal behaviour model trained from historical data) to identify the abnormal behavior. This baseline is specified by the (network) administrator. The network traffic and host activity/processes are usually monitored in anomaly-based detection. The system detects abnormal system behavior, identifying suspicious activities, looking up the ports and protocols used, and indicate any malicious activities to alert the administrators (Araújo and Abdelouahab, 2012; Oktay and Sahingoz, 2013).

A key focus is on incorporating anomaly detection techniques within IDS to identify novel and unknown threats that evade signature-based methods.

Artificial neural network (ANN) based IDS. In this type of IDS the programs' internal states and generally its behavior are monitored in order to defeat the infection. It utilizes ANN (Artificial Neural Network) as a pattern recognition technique. In particular, after monitoring the normal behavior of programs, it is used to train an ANN (Oktay and Sahingoz, 2013). The ANN model produced can then detect abnormal program behaviour that could signify its potential infection or unauthorized modification. Artificial Neural Networks (ANNs) are a form of machine learning algorithm inspired by the behavior of biological neurons located in the brain and the central nervous system.

Inputs to the ANN are typically fed to the artificial neurons in one or more hidden layers, where they are weighted and processed to decide the output to the next layer. ANNs make use of a "learning rule" (often gradient descent based back-propagation of errors) that allows the set of weights and biases for the hidden layer and output layer neurons to be adaptively tuned.

This makes them particularly useful in fields, such as decision support for concealed weapons detection, prediction and classification of Internet traffic, and signature verification, where their ability to adapt to high dimensional datasets overcomes many of the difficulties in model building associated with conventional classification techniques, such as decision trees and k-nearest neighbor algorithms. The intelligent intrusion detection system (Shenfield A., et al, 2018[90]) significantly improves upon the performance of signature-based detection methods by utilizing an artificial neural network classifier for the identification of intrusion patterns in network traffic.

[90] Alex Shenfield, David Day, Aladdin Ayesh, 2018, "*Intelligent intrusion detection systems using artificial neural* networks", ICT Express Volume 4, Issue 2, June 2018, Pages 95-99. https://doi.org/10.1016/j.icte.2018.04.003

ANN based IDS differs from Anomaly-based IDS, due to that the latter identifies deviations from normal behavior rather than relying on predefined attack signatures. It flags any unusual activity that doesn't fit known traffic patterns, as opposed to ANN based IDS, which uses deep learning to recognize attack patterns.

Fuzzy logic-based IDS. IDS which are increasingly a key part of system defense are used to identify abnormal activities in a computer system. In general, the traditional intrusion detection relies on the extensive knowledge of security experts on their familiarity with the computer system to be protected. To reduce this dependence, various data mining and machine learning techniques have been used. But this method (Oktay and Sahingoz, 2013) is based on rule-based, fuzzy-logic and data mining techniques so as to detect the intrusive behavior of network traffic. The current approach differs from Anomaly-based IDS in their decision-making approach, accuracy, and adaptability. Fuzzy logic-based IDS assess uncertain or imprecise attack conditions and handles uncertainty using degrees of truth (0-1) instead of strict rules.

The fuzzy logic-based system can be able to detect an intrusion behavior of the networks since the rule base contains a better set of rules. The strategy for generation the better set of fuzzy rules, is to obtain from the definite rules from the KDD CUP 99 Intrusion Detection Dataset, frequent items (mining of single length frequent items, identification of suitable attributes for rule generation, rule generation, rule filtering, generation fuzzy rules).

The KDD CUP 99 Intrusion Detection Dataset is one of the most widely used benchmark datasets for evaluating Intrusion Detection Systems (IDS), including mainly fuzzy logic-based IDS approaches (Shanmugavadivu, R., et. al., 2012[91]). It contains a large volume of simulated network traffic, categorized into normal activities and attack types such as Denial of Service (DoS). Fuzzy logic-based IDS systems utilize this dataset to train models that can detect and classify cyber threats with greater flexibility. Unlike traditional rule-based or signature-based IDS, which rely on strict thresholds, fuzzy logic can handle uncertainty and partial truths in network traffic analysis. It assigns degrees of suspicion rather than making rigid binary decisions (attack or no attack), making it more effective in identifying ambiguous or evolving attack behaviors. By applying fuzzy logic to the KDD CUP 99 dataset, IDS can adapt dynamically to variations in network activity, reducing false positives while maintaining high detection rates. The dataset's vast number of features allows fuzzy-based IDS to define flexible rules for attack detection, improving overall network security. This approach is particularly valuable in real-world cybersecurity, where threats are constantly changing, and traditional detection methods struggle to keep up.

Since there is no ideal solution to avoid intrusions from events, it is very significant to detect them at the initial moment of happening and take necessary actions for reducing the likely damage. One approach to handling suspicious behaviors inside a network is an intrusion detection system (IDS). One fuzzy rule is by learning for effective intrusion detection using data mining techniques. The fuzzy rules generated from the proposed strategy can be able to provide better classification rate in detecting the intrusion behavior. Fuzzy logic-based has lower false positive rate than the

[91] Shanmugavadivu R, Dr.N.Nagarajan, 2012, "*Network Intrusion Detection System using Fuzzy Logic*", Indian Journal of Computer Science and Engineering, February 2012, Vol. 2, No 1, pp 101-111

anomaly-based, since fuzzy logic can tolerate slight variations, whereas anomaly-based approach is higher due to strict normal behavior definitions.

5.1.2.3 Intrusion detection system types

There are some major types of IDS which are used in CC.

Network based intrusion detection system (NIDS)

NIDS observes, examines and analyzes the traffic to and from all the devices on the network. It can scan all inbound and outbound traffic and detect various types of instances. Once the attacks have been detected from this method, the active system immediately takes necessary actions and informs the administrators (Araújo and Abdelouahab, 2012; Oktay and Sahingoz, 2013).

The cloud provider's responsibility is to manage and deploy NIDS in the cloud environment. These types of NIDS can be useful to also detect some kinds of attacks in VMs and hypervisors. They can be placed in VPCs, edges of the cloud with the outside world, in front of load balancers, in between different (or same) virtual networks to monitor VM traffic, in provider network boundaries (cover traffic between regions and zones), management systems), etc. Moreover, the network traffic within and outside of the cloud environment cannot be decrypted to analyze it and this is a challenge for the NIDS.

Host based intrusion detection systems (HIDS)

This method runs on hosts or devices to analyze and observe them. HIDS monitor several entities (inbound and outbound packets, logs, processes, etc) in order to identify prospective attacks and attempts of unauthorized access in the supervised machines. Once suspicious activities are detected, the HIDS alerts the users or administrators.

In a cloud environment, an HIDS can be applied either on VMs or host machines. The HIDS deployed on a VM can be monitored and managed by customers. However, the HIDS deployed on a physical machine is monitored and administered by the cloud provider (Araújo and Abdelouahab, 2012).

Hypervisor-based intrusion detection system

A hypervisor or VMM implements the abstract layer between the underlying host and the guest operating system. This layer provides security, protecting the kernel of the host, so that the kernel becomes free of vulnerabilities. In this way, the HIDS can be placed in the hypervisor so as to examine related virtualization metrics coming from the hypervisor and detect any potential security issue.

This detection method allows monitoring and analyzing the communication between VMs, hypervisors and inside the virtual networks. This approach can be effective in a cloud environment. However, as this method is new, it will take some time to mature and be fully adopted by all cloud providers (Bharadwaja et al., 2011; Nikolai and Wang, 2014[92]).

[92] Bhadauria, R., et al., 2011. “A survey on security issues in cloud computing”. arXiv preprint arXiv:1109.5388, 2011.
Nikolai, J., Wang, Y., 2014. “*Hypervisor-based cloud intrusion detection system*”. In: Proceedings of the Computing, Networking and Communications (ICNC), 2014 International Conference on. 2014. IEEE.

5.1.2.4 Intrusion detection in cloud computing service models

Software as a Service: in SaaS users just consume SaaS services that are fully administered by their providers – this also covers the security part. Therefore, it is also the responsibility of the SaaS CSPs to deploy IDSes for security purposes.

Platform as a Service: in PaaS, IDS is merely deployed outside of the applications; thus, most of the IDS in PaaS are deployed by the CSP. The users can configure their applications and platform to place logs onto a central location to enable monitoring and alerting of the IDS – in some cases, they might have restricted access to such logs while they might be alerted when security issues take place.

Infrastructure as a Service: IaaS is considered a more flexible model for IDS as an IDS can be deployed in different IaaS parts, such as:

- The *VM*: deploying IDSes in the virtual machines can allow monitoring the system and subsequently alerting and detecting any unusual behavior of the VMs. In this way, the customers themselves alert the administrators about any unusual behavior. Alternatively, the administrator can do the detection and response about unusual behavior, given that VM comes with the IDS built-in.
- The *Hypervisor or host system*: As the hypervisor provides an abstraction layer, it allows the IDS to monitor the hypervisor itself and anything traveling between hypervisor and virtual machines. Thus, the IDS acts in this case like a central location for intrusion detection in the IaaS environment.
- *Virtual Network*: so as to be able to monitor the network traffic between VMs and host machines and between VMs themselves. In a virtual network, the IDS could be placed, for instance, at the hypervisor level (so same solution as previous one) or at the virtual switch that handles intra-host network traffic.

Finally, a *Hybrid Intrusion Detection System* (IDS) enhances cloud privacy by combining *signature-based* and *anomaly-based detection* techniques. In cloud environments, traditional IDS struggle due to dynamic workloads and evolving threats. A hybrid IDS leverages signature-based detection for known attacks and anomaly-based detection to identify unknown or zero-day threats. Machine learning and fuzzy logic further improve accuracy by reducing false positives. This approach ensures real-time monitoring, adaptive threat detection, and enhanced data privacy in the cloud. By integrating multiple detection methods, hybrid IDS strengthens cloud security, preventing unauthorized access, data leaks, and cyberattacks while maintaining system performance.

Below, in Table 7 we present our proposed countermeasures for the IDS category, with their best use, strengths and weaknesses. The proposed countermeasures for the IDS category, have been resulted from our research on the literature (see also Ch. 3, Methodology review), which we consider them as the most used.

Table 7: IDS countermeasures category & their best use cases

IDS Type	Best For	Strengths	Weaknesses	Value
Network-Based IDS (NIDS)	Monitoring network traffic for anomalies	Detects large-scale attacks & lateral movement	Needs high-performance infrastructure	*High* (Monitors cloud network traffic for suspicious activity in real time)
Host-Based IDS (HIDS)	Protecting individual endpoints & servers	Detects malware, unauthorized changes, file integrity issues	High overhead, limited to specific hosts	*High* (Monitors specific cloud hosts for malicious activity and policy violations)
Cloud IDS	Cloud-native security monitoring	Scalable, integrates with cloud workloads	Cloud provider-specific, may need third-party support	*High* (Cloud-native IDS that monitors cloud infrastructure for malicious activity)
Signature-Based IDS	Identifying known attacks & malware	Low false positives, fast detection	Fails against zero-day & evolving threats	*High* (Detects known attack patterns using a database of signatures)
Anomaly-Based IDS (AI/ML-driven)	Detecting unknown attacks & insider threats	Identifies zero-day threats & novel attack patterns	High false positives if not properly tuned	*High* (Identifies deviations from normal network behavior to detect potential intrusions.)
Hybrid IDS (Combining Signature + Anomaly Detection)	Comprehensive security monitoring	Balances accuracy & real-time detection	Requires more resources & tuning	*High* (Integrates host and network-based monitoring for comprehensive threat detection)
Basic Host-Based IDS	Small-scale environments or isolated servers with limited access	Easy to deploy on individual servers	Limited to single hosts, cannot monitor cloud-wide traffic	*Low* (It can catch some attacks, its low effectiveness arises from its inability to monitor network traffic across the

				cloud infrastructure)
Intrusion Prevention System (IPS)	Actively blocking threats in real time	Prevent attacks before they reach critical systems	Higher risk of false positives disrupting operations	*Moderate* (Uses external threat feeds to enhance detection capabilities)

5.1.3 DLP (Data Leakage Prevention)

DLPs differ from conventional security controls such as firewalls, VPNs and IDSs in terms of dedication and proactivity (see also at 2.2.6 Data Loss-Leakage and DLPs definition in Chapter 2). Conventional security controls have less dedication towards the actual content of the data. They might block users' access to data for the sake of sensitive data protection in the case of firewalls, or simply encrypt all the traffic, as in the case of VPNs, which might include both sensitive and non-sensitive data.

5.1.3.1 Techniques

Between a technique and an approach for data privacy protection, there is a difference. A technique is a technical or practical method or skill for completing a specific task while an approach maps to practical actions or mechanisms intended to deal with a problem or situation related to the privacy (e.g., in the cloud). In this subsection, we enumerate different techniques used to preserve privacy in cloud computing.

(I) Encryption.

Encryption is the most used and recognized security technique to ensure data confidentiality. It is about using a cryptographic solution to encrypt data stored in the cloud (or at transit). This process could be done either by the data owner or by the CSP. Cryptography in the cloud still represents an open issue since it requires the management of the cryptographic keys. The main question to ask is should we use symmetric key cryptography requiring the same key for encrypting and decrypting (it can be considered as a complex problem especially when different actors are involved) or should we use asymmetric cryptography based on a couple of private and public keys. Some encryption schemes that include symmetric and asymmetric encryption methods emerged in the cloud, but if the system involves multiple actors, and secure key exchange is a concern, symmetric key cryptography alone might not be ideal. A hybrid model or a fully asymmetric approach (e.g., using public-key infrastructure, or PKI) could be a better choice. On the other hand, if the environment allows secure pre-distribution of keys (e.g., within a closed system or single organization), symmetric encryption could suffice.

Some encryption algorithms are: *Triple Data Encryption Standard*, *Advanced Encryption Standard*, *RSA Security*, *Blowfish*, and *Twofish*. Also, started in 2016 as the NIST's Post-Quantum Cryptography Standardization project, four algorithms have been winnowed down from 69 submissions by cryptography experts in dozens of countries: *CRYSTALS-Kyber* (FIPS 203),

designed for general encryption purposes, such as creating websites, *CRYSTALS-Dilithium* (FIPS 204), designed to protect the digital signatures used when signing documents remotely, *SPHINCS+* (FIPS 205) is also designed for digital signatures and *FALCON* is also designed for digital signatures and is slated to receive its own draft FIPS in 2024[93]. These four algorithms were released by NIST as a final set of encryption tools designed to withstand the attack of a quantum computer. These post-quantum encryption standards-algorithms secure a wide range of electronic information, from confidential email messages to e-commerce transactions that propel the modern economy. NIST is encouraging computer system administrators to begin transitioning to the new standards as soon as possible.

(II) Processing encrypted data

To overcome the limitations of data encryption, an encrypted data processing technique was proposed to allow performing computation on encrypted data. Therefore, the cloud actor does not need to decrypt data for query execution and can execute queries directly on the encrypted data. Multiparty computation, such as Yao's secure two-party protocol (Yao C., 1982[94])l, enables a data user and a data owner to cooperate to calculate a function over data without data disclosure. Must be noted, that querying is different from processing. In querying, there is no data modification but in processing someone might modify the data. Homomorphic encryption, such as Gentry's[95] (2009) fully homomorphic encryption scheme, provides a *general* way of calculating a function on encrypted data in the cloud. It generates an encrypted result which, when decrypted, represents the result of the same function if performed on the respective plain-text form of the data.

(III) Obfuscation

Data obfuscation or data masking refers to the process that enables concealing sensitive information by replacing it with realist-looking values based on secret masking rules. These values look similar but are significantly different (unrelated) from the real data. Data obfuscation offers three main techniques that fall under the category of privacy preserving techniques: (a) *Data Randomization* (a technique used to protect sensitive information by introducing randomness into the data. It works by modifying the original data in such a way that individual data points are obscured, making it difficult for attackers or unauthorized parties to infer private information, while still allowing for meaningful statistical or analytical computations on the randomized data*),* (b) *Data Swapping* (swaps entries within a single field in a record set so that the individual record entries are unmatched or intelligently swaps the pixels of an image) and (c) *Anonymization* (falls within obfuscation techniques that include removing personally identifiable information - PII - from a data record. See also 5.1.4.-IV).

(IV) Sticky policy

Sticky policy is an advanced technique coming from the need to gather data along with the user preferences regarding privacy. This technique allows sticking individual privacy circumstances and preferences (formatted as policies) directly to personal data while moving them across the cloud.

---

[93] https://www.nist.gov/news-events/news/2024/08/nist-releases-first-3-finalized-post-quantum-encryption-standards

[94] Yao's Protocol: Andrew C.-C. Yao: *"Protocols for Secure Computations" Proceedings of the 23rd Annual Symposium on Foundations of Computer Science (FOCS), 1982.* This is the foundational paper where Yao first proposed the idea of secure multiparty computation (SMC) and the garbled circuits protocol for two-party computation. DOI: 10.1109/SFCS.1982.38

[95] Gentry C (2009) May, "*Fully homomorphic encryption using ideal lattices*". STOC 9:169–178

Hence, data processing is permitted strictly when the attached policies are respected. The effectiveness of the sticky policies technique strongly relies on the effectiveness of their enforcement in the cloud. Therefore, it is crucial to introduce powerful means for policy enforcement in the cloud.

(V) Data segmentation

Data segmentation is an additional solution to ensure data confidentiality when relying on external cloud providers for storing and processing data. This technique is based on the fact that private data is not only sensitive information; the associations between data are also sensitive. Thus, privacy can be guaranteed by storing different chunks of the data in separate non-linkable fragments.

(VI) Trusted third party mediator (TTPM)

Trusted third party would act as a mediator (TTPM) between cloud actors to ensure the policy enforcement and carry out the audit or even with management. This technique is not new (Wang B. et, al., 2013[96]), it was adopted to build online customer trust when using e-commerce applications. A trusted third-party (TTP) mediator is a neutral entity or organization that facilitates secure interactions, communication, or transactions between two or more parties. The mediator is trusted by all involved parties to ensure fairness, integrity, and privacy. TTPs are commonly used in scenarios where the direct exchange of sensitive information, resources, or trust is impractical or insecure.

(VII) Trusted platform module

Trusted Platform Module (TPM) is a tamper-resistant hardware component developed by the Trusted Computing Group (TCG). This hardware component acts as the root of trust. In fact, TPM provides the ability to perform security-related actions, such as identification and authentication, encryption and decryption, key generation, signature creation, hash generation, integrity checking of software, nonvolatile storage of secrets, etc.

In case there is s need for strengthening Administration for Data Leakage prevention, we can confront it, with:

- *Privileged Access Management (PAM)*: Implementing PAM tools like CyberArk, BeyondTrust, or HashiCorp Vault[97] to secure and monitor privileged accounts.

- *Separation of Duties*: Ensuring that no single administrator has unrestricted access to all aspects of the cloud. Divide responsibilities across different roles (e.g., security admin, resource admin), and

- *Automation*: Automate tenant onboarding, identity provisioning, and lifecycle management using orchestration tools like Ansible, Terraform, or CloudStack[98].

[96] Wang B., Chow S., Ki M., Hui L,, “Storing Shared Data on the Cloud via Security-Mediator”, Conference: Distributed Computing Systems (ICDCS), 2013 IEEE 33rd International Conference on Distributed Computing Systems, July 2013, DOI: 10.1109/ICDCS.2013.60

[97] https://www.cyberark.com/ , https://www.beyondtrust.com/ , https://www.vaultproject.io/

[98] https://docs.ansible.com/ , https://www.terraform.io/ , https://cloudstack.apache.org/

All the privacy techniques have their own advantages and disadvantages, which are summarised in Table 8. Encryption, for example, can partially address the challenges associated with malicious insiders by preventing them from obtaining private data in their plain-text format. However, encryption cannot be an eventual solution when data must be decrypted to be processed. Encryption generates some problems due to the additional complexity introduced in cloud computing models. Encryption introduces several challenges in cloud computing due to the added complexity it brings to data security, performance, and management. First, encryption and decryption operations consume significant computational resources, which can lead to performance degradation, especially when dealing with large volumes of data or real-time applications. Additionally, managing encryption keys securely becomes more complex in the cloud, as improper key management can lead to data breaches or loss of access to encrypted data. Furthermore, encrypted data is not easily usable for cloud-based operations like querying or analysis without decryption, limiting functionality. These challenges are compounded by the shared responsibility model in cloud environments, where both providers and customers must manage encryption tasks, increasing the risk of misconfiguration or vulnerabilities.

Table 8. Privacy techniques strengths and weaknesses (Ghorbel, A., et al, 2017)

Techniques	Strengths	Weakness
Encryption	Enables strong protection for data at rest	Data Must be decrypted to be processed
	Supports any type of data	Prevent indexing and searching
Processing encrypted data	Enables processing the encrypted data	Adds a high computation overhead
	Ensures data protection throughout its lifetime	Not always feasible in the cloud
	Supports any type of data	
Obfuscation	Enables to perform sufficient calculation accuracy masked data	Week protection than encryption
	Provides multiple obfuscation techniques	Not always feasible in the cloud
Sticky policy	Stucks data to policy	Adds a relatively high overhead
	Processing is permitted unless policies are respected	Its efficiency depend on the PEP
TPM	Provides a shielded location to protect user's data secrets	Cannot perform secure processing
		Presents a hardware solution
Segmentation	Enables protection for data at rest	The loss of a chunk of data leads to the loss of data in all
TTPM	Enables a trusted party to check for privacy compliance	Adds high-level computation due to the additional communication traffics

5.1.3.2 Approaches

As aforementioned, an approach differs from a technique in the sense that it is based on assumptions and theoretical ideas to solve a problem.

Data-centric approaches emphasize mechanisms and techniques to automate sensitive data protection anywhere in any distributed system. One of the relevant data-centric approaches is self-protecting data solutions. These solutions enable data to defend itself even though it is located in an untrusted environment. Squicciarini.A.C. et al. (2013[99]) introduced the Self-Controlling Objects (SCO). This object encapsulates sensitive data along with their policies and assures their protection by means of object-oriented programming techniques. The SCO policy concerns data owner policy and legal policy. Each time the access to the SCO protected content is attempted, its policy is evaluated according to the requester's credential and location.

[99] Squicciarini, A.C., Petracca, G., Bertino, E.: Adaptive data protection in distributed systems. In: Proceedings of the Third ACM Conference on Data and Application security and privacy pp. 365–376. ACM, February 2013

User-centric approaches. This kind of approach focuses on how to protect data from the data originator side by considering his/her preferences or legal texts. Some work under this category tries to define advanced tools to enable supporting properly and defining policies for data protection when data are processed in the cloud. Researchers (Papanikolaou N, Pearson S, Mont MC, 2014) have come up with a toolkit for automating compliance of cloud computing services. This toolkit allows the semantic annotation and natural-language processing of policy texts (regulation text and/or user’s preferences text) to generate machine-readable rules. The implementation of the proposed tool is incorporated in the EnCoRe Policy Enforcement Framework[100].

CSP-centric approaches. The CSP-centric approaches focus on mechanisms and frameworks to be deployed and integrated into the CSP infrastructure to assure data protection. The work presented by Yau, S. et. al. (2010), aims to protect data from the CSP. They propose three conditions under which the CSP may reveal private data: (i) CSP knows the storage location data, (ii) CSP has the privilege to access and collect data and (iii) CSP may understand data. The authors believe that if they can prevent the CSP from fulfilling the three conditions simultaneously, they can protect data. In this kind of approach, the user’s involvement is negligible (does not specify policies or preferences) compared to the CSP’s one.

Hybrid approaches. A hybrid approach is a solution that combines two or more approaches from the categories already presented. Some of the hybrid approaches rely on the sticky policy technique plus policy management components that ensure the sticky policy enforcement in a cloud environment.

5.1.3.3 Evaluation

We have studied the techniques adopted within the current approaches to cope with the privacy issues in cloud computing. We can note that all presented works endorse one or more of the techniques presented previously. Most used techniques for privacy preservation in the cloud are probably: encryption, the sticky policy and the privacy policy alone without being bound to data. This is due to the fact that these techniques look promising to address the privacy problems in the cloud with reasonable complexity, performance and economic overheads. On another side, some approaches come up with concepts and assumptions that are difficult to apply in a real context. To ensure better protection of his/her private data in the cloud, the data owner must make a precise specification of his/her policies and must consider eventually legal policies (e.g., GDPR). Privacy enforcement can be done at different levels including the high-level software, low-level software and hardware. The high-level software solutions introduce software that enable enforcement of high-level policies at application level. Data usage generates multiple data copies and derived/modified data in different locations across a computing system. Usage control and policy enforcement must be enforced for all the copies of data and the derived data through the overall system. Private data must be protected throughout its lifetime, whether it is in-transit, in-use or at-rest.

5.1.3.4 Deployment

[100] EnCoRe 2011. The EnCoRe project. http://www.encore-project.info/

Depending on the targeted data for protection, DLPs deployment can take many forms. For example, protecting “in use” data requires built-in software that acts like a DLP agent on endpoints. This agent is responsible for disabling or enabling access to applications that deal with confidential data. It is also responsible for blocking confidential data transfer through portable media, that is, CDs, USB drives and memory cards. Furthermore, it restricts copying, pasting and editing of confidential data as well as restricts making hard copies through printers. Finally, it audits all activities related to confidential data access. (Hackl and Hauer, 2009; McAfee, 2015). For “in transit” data, DLP appliances are normally used. They come with special processing capabilities to handle large amounts of data.

Below, in Table 9 we present our proposed countermeasures for the DLPs category, with their best use, strengths and weaknesses. The proposed countermeasures for the DLPs category, have been resulted from our research on the literature (see also Ch. 3, Methodology review), which we consider them as the most used

Table 9: DLPs countermeasures category & their best use cases

DLP Countermeasure	Best For	Strengths	Weaknesses	Value
Content Inspection (Regex, Fingerprinting, AI/ML-based Detection)	Preventing leaks of PII, financial, and proprietary data	Accurately detects sensitive data patterns	High false positives if rules are not well-configured	*High* (Analyzes file contents, metadata, and patterns to detect sensitive data before transfer)
Endpoint DLP (Agent-Based Monitoring, USB Control, File Encryption)	Preventing insider threats and unauthorized data transfers	Secures endpoints (laptops, desktops) from leaks	Can slow down endpoint performance	*High* (Prevents sensitive data from leaving managed endpoints via USB, clipboard, or network)
Network DLP (Traffic Analysis, TLS Inspection, Email Filtering)	Preventing data exfiltration via email, web, or cloud services	Detects & blocks unauthorized data movement	May require SSL/TLS decryption, raising privacy concerns	*High* (Monitors data flowing across cloud networks and blocks unauthorized transfers)
Cloud DLP (CASB, API-Based DLP, Cloud-Native DLP)	Protecting SaaS, IaaS, PaaS environments	Ensures cloud data security & compliance	Requires integration with multiple cloud providers	*High* (Prevents data leaks by enforcing policies across cloud services)
Encryption & Tokenization (AES-256, FPE, Hashing)	Securing data at rest, in transit, and in use	Protects data even if exfiltrated	Key management complexity	*High* (Protects data from unauthorized access by encrypting it during storage

				and transmission)
Access Control & Identity Management (RBAC, ABAC, Zero Trust)	Restricting access to sensitive data. RBAC ideal for Container vulnerabilities (	Reduces insider threats & unauthorized access	Poorly configured policies may lead to gaps	*Low* (Relies on human intervention, making enforcement slow and inconsistent)
AI-Driven Behavioral Analytics (UEBA, SIEM, Insider Threat Detection)	Identifying unusual data movements	Detects insider threats & anomalous behavior	Requires AI/ML tuning to reduce false alerts	*High* (Uses machine learning to analyze data movement and detect anomalous activities)
Basic Data Loss Prevention (DLP) with Keyword Matching	Non-sensitive environments with minimal risk of data leakage, or for monitoring common, simple patterns	Easy to deploy and configure	Vulnerable to evasion (synonyms, encoding, or obfuscation)	*Low* (can prevent obvious data leaks. However, it is a low-effectiveness countermeasure in cloud data protection because it is unable to detect complex or obfuscated data)
Backup & Data Redundancy (Immutable Backups, Versioning, Snapshots)	Preventing data loss due to ransomware & corruption	Ensures data recovery after attacks	Storage cost increases with redundancy	*High* (Stores copies of data across multiple geographic locations to ensure availability)

We analyze the above countermeasures, as follows:

- *Content Inspection* is a critical Data Loss Prevention (DLP) countermeasure that analyzes and monitors data to prevent unauthorized access, leaks, or breaches. It employs regular expressions (Regex) to detect sensitive patterns like credit card numbers, fingerprinting to track specific documents, and AI/ML-based detection for advanced content analysis. This approach ensures that confidential information in emails, cloud storage, and network traffic remains secure. AI-powered detection enhances accuracy by identifying contextual data patterns and reducing false positives. By combining these techniques, content inspection strengthens compliance, security, and regulatory adherence, protecting organizations from data exfiltration, insider threats, and cyberattacks.
- *Endpoint Data Loss Prevention (DLP)* is a crucial countermeasure that protects sensitive data on user devices. It uses agent-based monitoring to track file activities, USB control to restrict unauthorized data transfers, and file encryption to secure sensitive documents. By enforcing security policies on laptops, desktops, and mobile devices, endpoint DLP, prevents data leaks from insider

threats, malware, or unauthorized access. Real-time monitoring and policy enforcement ensure compliance with regulatory standards, reducing the risk of data breaches. This approach strengthens enterprise security by protecting sensitive information at its source, whether online or offline, within or outside the corporate network.

- *Network Data Loss Prevention (DLP)* safeguards sensitive information by monitoring and controlling data flows across a network. It uses traffic analysis to detect unauthorized data transfers, TLS inspection to analyze encrypted communications, and email filtering to prevent data leaks via emails. By inspecting incoming and outgoing network traffic, Network DLP helps organizations enforce security policies, detect anomalies, and prevent data exfiltration. This countermeasure is essential for ensuring regulatory compliance, protecting intellectual property, and mitigating insider threats or cyberattacks. With real-time monitoring and automated responses, Network DLP enhances data security across cloud, on-premises, and hybrid environments.
- *Cloud Data Loss Prevention (DLP)* protects sensitive data in cloud environments by monitoring and controlling data access and movement. It includes Cloud Access Security Brokers (CASB) to enforce security policies, API-based DLP for seamless integration with cloud services, and cloud-native DLP for built-in protection within platforms like AWS, Google Cloud, and Microsoft Azure. These solutions prevent unauthorized data sharing, misconfigurations, and insider threats while ensuring compliance with regulations like GDPR and HIPAA. By providing real-time visibility, encryption, and automated policy enforcement, Cloud DLP secures sensitive information in SaaS, IaaS, and PaaS environments, reducing the risk of cloud-based data breaches.
- *Encryption & Tokenization are essential Data Loss Prevention (DLP)* countermeasures that protect sensitive data by making it unreadable to unauthorized users. AES-256 encryption secures data at rest and in transit, ensuring strong protection against breaches. Format-Preserving Encryption (FPE) allows encrypted data to retain its original structure, making it useful for compliance-sensitive applications. Hashing converts data into fixed-length values for integrity verification. Tokenization replaces sensitive data with non-sensitive equivalents, reducing exposure risks. These techniques help organizations comply with GDPR, PCI-DSS, and HIPAA, safeguarding critical information against cyber threats, insider risks, and data leaks in on-premises and cloud environments.
- *Access Control & Identity Management are key Data Loss Prevention (DLP)* countermeasures that regulate data access based on user roles and policies. Role-Based Access Control (RBAC) grants permissions according to predefined job roles, while Attribute-Based Access Control (ABAC) enforces rules based on user attributes, environment, and resource sensitivity. Zero Trust Architecture ensures continuous verification, requiring authentication before granting access. These approaches limit unauthorized data access, prevent insider threats, and enhance compliance with regulations like GDPR and HIPAA. By enforcing privilege principles, organizations reduce the risk of data breaches, securing sensitive information across cloud, on-premise, and hybrid environments.
- *AI-Driven Behavioral Analytics is a powerful Data Loss Prevention (DLP)* countermeasure that detects anomalies and insider threats using User and Entity Behavior Analytics (UEBA), Security Information and Event Management (SIEM), and machine learning models. By analyzing user activities, access patterns, and system interactions, it identifies suspicious behavior, such as data exfiltration, privilege abuse, or unauthorized access. SIEM systems aggregate and correlate security events for real-time threat detection, while UEBA continuously monitors deviations from

normal behavior. This proactive approach enhances cybersecurity, regulatory compliance, and risk mitigation, reducing the likelihood of data breaches and improving overall enterprise security posture.

- *Backup & Data Redundancy is a crucial Data Loss Prevention (DLP)* countermeasure that ensures data availability and integrity in case of cyberattacks, accidental deletions, or system failures. Immutable backups protect data from ransomware by preventing unauthorized modifications or deletions. Versioning maintains historical copies of files, allowing recovery from accidental changes. Snapshots capture the system’s state at specific points, enabling quick restoration after data loss incidents. These techniques strengthen disaster recovery, business continuity, and regulatory compliance by providing reliable data protection. Implementing secure, redundant storage strategies minimizes downtime and mitigates risks associated with data breaches or corruption.

5.1.4 Data storage and PasS (Privacy as a Service)

PasS (Privacy as a Service) is in general a set of security protocols for ensuring the privacy and legal compliance of customer data in cloud computing architectures (Itani W., at al. 2009). PasS allows for the secure storage and processing of users’ confidential data by leveraging the tamper-proof capabilities of cryptographic co-processors. Using tamper-proof facilities provides a secure execution domain in the computing cloud that is physically and logically protected from unauthorized access.

This is achieved by implementing user-configurable software protection and data privacy mechanisms. Moreover, PasS provides a privacy feedback process which informs users of the different privacy operations applied on their data and makes them aware of any potential risks that may jeopardize the confidentiality of these data. To the best of our knowledge, PasS is the first practical cloud computing privacy solution that utilizes previous research on cryptographic coprocessors to solve the problem of securely processing sensitive data in cloud computing infrastructures. PasS central design goal is to maximize users’ control in managing the various aspects related to the privacy of sensitive data. This is achieved by implementing user-configurable software protection and data privacy categorization mechanisms. Moreover, PasS provides a privacy feedback process which informs users of the different privacy operations applied on their data. We can also refer to some data storage privacy protocols: *Data and Software Transfer Protocol*[101], *Software Execution and Data Processing Protocol*[102], and *Privacy Feedback Protocol*[103].

[101] The *Data and Software Transfer Protocol* ensures secure and efficient data storage by providing guidelines for encrypting and transferring sensitive information. It maintains data integrity, prevents unauthorized access, and supports cloud-based storage solutions, enhancing data confidentiality and compliance with regulatory standards, such as GDPR and HIPAA, across various platforms.

[102] The *Software Execution and Data Processing Protocol* ensures secure management of data storage during software operations. It defines encryption, access control, and integrity verification methods to protect stored data. By enforcing strict security standards, this protocol enhances data privacy and ensures compliance with regulatory requirements, safeguarding sensitive information in storage environments.

[103] The *Privacy Feedback Protocol* monitors and manages data storage to ensure compliance with privacy policies. It enables real-time feedback on data handling practices, ensuring that sensitive information is securely stored and processed. This protocol helps maintain data confidentiality, prevent unauthorized access, and align with privacy regulations like GDPR.

Despite all the advantages delivered by CC, several challenges are hindering the migration of customer software and data into the cloud. On top of the list are the security and *privacy concerns* arising from the *storage and processing* of sensitive data on remote machines that are not owned or even managed by the customers themselves. With CC, what all the customers can see is a virtual infrastructure built on top of possibly non-trusted physical hardware or operating environments (unless the CSP has monitoring and observability mechanisms that enable him to sneak on user data). Data privacy should be provided to cloud customers as a service with minimal additional cost. Moreover, we believe that the cloud privacy model should be configurable and user centric.

Locating data can help users to increase their trust in the cloud. Cloud storage provides the transparent storage service for users, which can decrease the cloud’s complexity. However, it also decreases the control ability of users on data storage.

Cloud data storage is integral to the operations of Privacy as a Service (PasS), which aims at protecting sensitive information while ensuring compliance with privacy regulations. The complexity of cloud ecosystems introduces unique privacy challenges, necessitating robust countermeasures to safeguard data from threats like unauthorized access, breaches, and regulatory non-compliance. Below is an analysis of key countermeasures for cloud privacy in PasS environments, focusing on their effectiveness, challenges, and implementation considerations.

Here are the main countermeasures and prevention techniques:

(I) Data Encryption

Effectiveness: Encryption is a cornerstone of cloud privacy, protecting data at rest and in transit. Algorithms like AES-256[104] ensure robust protection, while TLS secures communication channels. Homomorphic encryption[105] offers an advanced option, allowing computations on encrypted data without decryption. Data encryption is a prevention technique used also for the Data Leakage Prevention scheme (category).

Challenges: Effective encryption requires meticulous key management. Mismanagement of encryption keys can lead to data inaccessibility or compromise. Homomorphic encryption, though promising, incurs significant computational costs, making it less practical for real-time applications. In addition, it does not support all sorts of processing over data.

Implementation: PasS providers must integrate encryption into their storage services while offering user-friendly key management systems, such as Hardware Security Modules (HSMs).

(II) Access Control and Authentication

Effectiveness: Granular access control models and Multifactor Authentication (MFA) mitigate risks of unauthorized access. The Zero Trust Architecture enhances security by requiring continuous verification of users and devices. A Zero Trust architecture (ZTA) is an enterprise

[104] The Advanced Encryption Standard (AES) is the algorithm trusted as the standard by the U.S. Government and numerous organizations and is also found in Arcserve Unified Data Protection (UDP) software. Although it is highly efficient in 128-bit form, AES also uses keys of 192 and 256 bits for heavy-duty encryption purposes. AES is largely considered impervious to all attacks, except for brute force, which attempts to decipher messages using all possible combinations in the 128, 192, or 256-bit cipher.

[105] Homomorphic encryption is a cryptographic technique that allows computations to be performed on encrypted data without requiring decryption. Current encryption technology is incredibly valuable. It underpins the Internet, lies at the heart of Web3, and empowers individuals to protect their personal information.

cybersecurity architecture that is based on zero trust principles and designed to prevent data breaches and limit internal lateral movement.

Challenges: Implementing fine-grained controls across large user bases can become administratively complex. MFA might hinder usability, especially if users face frequent authentication prompts.

Implementation: Privacy-as-a-Service platforms can use AI-driven behavioral analytics to streamline authentication processes while maintaining security.

(III) Data Masking and Tokenization

Effectiveness: These methods safeguard sensitive information while preserving its usability for specific operations. Tokenization is particularly effective for compliance with regulations like PCI DSS, while masking is useful in non-production environments (see also 5.2.2.1-III). Tokenization in Cloud Data Security is a process of replacing sensitive data, such as personal identifiers or payment information, with non-sensitive tokens that can be securely stored and transmitted. Tokenization is widely used in cloud environments to enhance data security while minimizing the risk of data breaches.

Challenges: Masked or tokenized data may not be suitable for all workflows, particularly where real-time analytics or advanced processing is required.

Implementation: Tokenization can be integrated into PaaS environments for sensitive data, such as PII or payment details. Dynamic masking techniques can adapt based on user roles or context.

(IV) Anonymization and Pseudonymization

Effectiveness: These techniques reduce the risk of re-identification, ensuring compliance with privacy laws, such as GDPR. Anonymization is ideal for static datasets, while pseudonymization supports scenarios requiring reversible data transformations. Pseudonymization is a data management and de-identification procedure by which personally identifiable information fields within a data record are replaced by one or more artificial identifiers, or pseudonyms. A single pseudonym for each replaced field or collection of replaced fields makes the data record less identifiable while remaining suitable for data analysis and data processing. Pseudonymization supports scenarios requiring reversible data transformations by allowing sensitive information to be replaced with pseudonyms that can be reverted to the original data when necessary. This reversibility is critical in use cases where organizations need to protect sensitive data but still access it for specific operational, legal, or compliance purposes (financial services and banking, health care data management, insurance claims processing, etc)..

Data anonymization has been defined as a "process by which personal data is altered in such a way that a data subject can no longer be identified directly or indirectly, either by the data controller alone or in collaboration with any other party." Data anonymization may enable the transfer of information across a boundary, such as between two departments within an agency or between two agencies, while reducing the risk of unintended disclosure, and in certain environments in a manner that enables evaluation and analytics post-anonymization.

Challenges: Effective anonymization is more complicated, as improper techniques may leave data vulnerable to re-identification attacks. Pseudonymization, while reversible, requires stringent key management.

Implementation: PaaS platforms must implement robust and scalable algorithms for k-anonymity, l-diversity, and differential privacy, especially for data analytics services.

(V) Secure Backup and Recovery

Effectiveness: Regularly encrypted backups stored in geographically dispersed locations protect against ransomware and natural disasters. Snapshots provide rapid recovery, minimizing downtime.

Challenges: Backup solutions may face (and cause) performance bottlenecks during high-volume operations. Over-reliance on the cloud provider’s infrastructure may lead to vendor lock-in, as the performance of high-volume operations can be also influenced.

Implementation: Implement immutable backups and periodically test recovery procedures. Use multi-cloud strategies to distribute risk.

(VI) Privacy Auditing and Monitoring

Effectiveness: Continuous monitoring and audit trails enhance visibility, enabling detection of anomalous behavior and ensuring regulatory compliance.

Challenges: Generating actionable insights from large volumes of log data requires advanced analytics and can strain resources. False positives in monitoring systems may lead to alert fatigue.

Implementation: Employ AI-driven Security Information and Event Management (SIEM) tools to automate threat detection and response, as they enable less resources and minimise false positives.

(VII) Data Minimization

Effectiveness: By reducing the amount of sensitive data stored, data minimization limits the impact of breaches and aligns with principles of privacy regulations like GDPR. This technique applies from the cloud customer side, as the one who can decide about which data should be stored is, actually, the customer.

Challenges: Determining which data is essential for operations can be difficult, particularly in data-driven industries.

Implementation: Implement automated data lifecycle management tools to enforce retention policies.

(VIII) Regulatory Compliance and Sovereignty

Effectiveness: Adhering to laws such as GDPR, HIPAA, and CCPA ensures legal protection and builds user trust. Data sovereignty measures, like storing data in specific geographic regions,

enhance compliance. The regulatory compliance, here, is referenced as an indirect technique for data storage, as the compliance demands to undertake various data storage protection measures.

Challenges: Balancing cross-border data sharing with compliance requirements can be complex, especially for multi-national organizations.

Implementation: Privacy-as-a-Service providers can offer compliance dashboards and configurable data residency settings.

PasS providers must implement a multi-layered privacy approach to address the complexity of cloud data storage threats. No single countermeasure is sufficient, as each addresses a specific aspect of privacy:

Encryption ensures baseline data security, but its effectiveness relies on strong key management.

Access controls and authentication mitigate internal and external threats but require fine-tuning for user convenience.

Data masking and tokenization offer lightweight alternatives for scenarios where encryption might hinder performance.

Anonymization and pseudonymization balance privacy and data utility but require rigorous implementation to avoid re-identification risks.

Regular backups and monitoring ensure resilience and accountability, essential for mitigating breaches and recovery.

By combining these strategies, organizations can achieve robust privacy protections while optimizing usability and operational efficiency. The integration of AI-driven tools and privacy-by-design principles will further enhance the scalability and effectiveness of PaaS platforms.

Below, in Table 10 we present our proposed countermeasures for the Data storage (and PasS) category, with their best use, strengths and weaknesses. The proposed countermeasures for the Data storage (and PasS) category, have been resulted from our research on the literature (see also Ch. 3, Methodology review), which we consider them as the most used. It is important to note that, for every countermeasure in this category, we present its positive and negative characteristics of its impact.

Table 10: Data storage (and PasS) countermeasures category & their best use cases

Countermeasure	Best For	Strengths	Weaknesses	Value
Encryption (AES-256, Homomorphic, FHE, TEE)	Protecting data at rest, in transit, & in use	Strong data confidentiality	Performance overhead	*High* (Provides homomorphic encryption to process encrypted data in a secure environment, ensuring privacy while minimizing performance overhead)

Tokenization & Data Masking	Protecting sensitive data (PII, PCI, PHI)	Irreversible & secures sensitive fields	Storage complexity	*High* (Replaces sensitive data with tokens, minimizing exposure while maintaining usability)
Immutable Storage & WORM (Write-Once-Read-Many)	Preventing data tampering & ransomware	Ensures data integrity	Cannot be modified once written	*High* (Prevents modifications or deletions of stored data, protecting against ransomware and insider threats)
Secure Backup & Disaster Recovery (BaaS, DRaaS, Snapshots)	Data loss prevention & high availability	Redundancy & quick recovery	Requires storage & cost planning	High (Provides comprehensive disaster recovery services using cloud infrastructure, enabling quick restoration of data and applications in case of a disaster, ensuring minimal downtime and data loss)
Multi-Cloud & Hybrid Storage Redundancy	Avoiding vendor lock-in & resilience	Increased availability & failover	Complex management	*High* (Distributes data across multiple cloud platforms to reduce risks associated with vendor lock-in and improve availability)
Confidential Computing (TEE, SGX, SEV, Nitro Enclaves)	Securing data in use (computation security)	Protects sensitive computations	Limited adoption & hardware dependency	*High* (Encrypts data even while being processed, preventing unauthorized access during computations)
Data Loss Prevention (DLP) for Cloud Storage & Monitoring	Preventing unauthorized data transfers	Tracks & blocks leaks in real-time	Needs tuning to reduce false positives	*Moderate* (Identifies and prevents sensitive data leaks but may require additional tuning for effectiveness.)
Blockchain for Data Integrity & Auditability	Ensuring tamper-proof logs & audit trails	Prevents unauthorized modifications	Storage overhead & complexity	*High* (Ensures data integrity by providing tamper-proof logs and

				decentralized verification)
Local Storage with Manual Backups	Small-scale environments or non-critical applications where data security risks are minimal	Simple to set up and use and low upfront cost	- Prone to human error (e.g., missed backups) - Limited scalability - Backup data might be insecure (lacking encryption)	*Low* (Using local storage with manual backups for cloud data protection is a low-effectiveness countermeasure due to its lack of automation, scalability, and protection)
Secure Erasure & Data Shredding	Ensuring secure deletion of sensitive data	Compliance with GDPR, HIPAA, NIST 800-88	Must be carefully implemented	*High* (Overwrites data before deletion to ensure it is unrecoverable, preventing unauthorized recovery)

5.1.5 Data protection mechanisms as countermeasures

5.1.5.1 Cryptography-Oriented Privacy Measures.

Cloud storage preserves data as if it is stacked in a locker, where the cloud storage acts as an access control to this locker. The data are encrypted, but usually the cloud server has the decryption key which manages the rights for each user to access the data. While this simplifies access management, it introduces potential security and privacy concerns. There are advantages to server-managed decryption keys but also challenges and risks that must be mitigated. Implementing client-side encryption, trusted key management systems, and robust security controls can enhance data protection in cloud environments. For sensitive or regulated data, adopting models like end-to-end encryption or zero-knowledge encryption is crucial to ensuring compliance and safeguarding user privacy. This is a critical problem in the case of private sensitive data, such as administrative documents or, more generally, personal data. This is quite tricky in the case of confidential documents possessed by a business enterprise, i.e., shared between the collaborators or with trading partners. In fact, this problem can be effortlessly solved by merely encrypting the data before sending it to the cloud/safe. Different architectures were proposed to resolve this problem which worked over two policies Trust-evaluation and predicate encryption.

5.1.5.2 Probability-Oriented Privacy Measures

In support of privacy by probability one best approach is the following:

Effective Solution for Data Leakage. Papadimitriou et. al (2010) deliberate the following problem: ‘‘A data distributor gives sensitive data to a set of allegedly trusted agents (third parties). Some of the leaked data is found in an unauthorized place (e.g., on somebody’s laptop or on the web). The distributor must weigh the outlook that the leaked data came from one or more agents, as opposed to having been independently gathered by other unauthorized means’’. In concern to

agents, the authors have proposed data allocation strategies that improve the probability of identifying leakages. Such methods do not depend on alterations of the released data (e.g., watermarks). In some cases, authors also inject ''realistic but fake'' data records to further improve the chances of detecting leakage and identifying the guilty agent or party.

5.1.5.3 Probabilistic Hybrid Logics

Hsu et al. (2018) proposed a combination of basic hybrid logic and quantitative uncertainty logic with a satisfaction operator. The logic is expressive and flexible enough to represent many existing privacy criteria, such as k-anonymity, logical safety, l-diversity, t-closeness, and d-disclosure. The main contribution of utilizing logic is twofold. On the one hand, the uniformity of the framework explicates the common principle behind a variety of privacy requirements and highlights their differences. For example, the difference between syntactic and semantic privacy criteria is easily observed by using logical specifications. On the other hand, the generality of the framework extends the scope of privacy specifications.

5.1.5.4 Utility Preserving Data Clustering

Nayahi and Kavitha (2016) proposed a KNN(G,S)[106] clustering algorithm to achieve anonymized clusters, each with uniform distribution of sensitive values. The (G,S) clustering algorithm governs the single best neighbor of each cluster and allocates from the neighbor one instance (of the neighbor cluster) at a time to the existing cluster. The authors have modified the algorithm to overcome the skew in the sensitive value distribution of the resultant clusters using the K-nearest neighbor technique.

5.1.5.5 Anonymization-Oriented Privacy Measures

This category of privacy perseverance comprises subset generations of well-known methods, anonymization methods. Based on secured multiparty computation, collaborative privacy-preserving data mining sustains high computational cost and communication. Data anonymization is an encouraging technique in the field of '*privacy preserving data mining*'[107] to protect the data against identity disclosure (Lindell Y. et. al, 2000). The 'privacy preserving data mining' model based on decision tree learning demonstrates that preserving data mining from large databases, can be efficiently achieved through secure multi-party computation. By utilizing cryptographic protocols, it is significantly reduced communication overhead, making secure data mining practical even for large and complex datasets, with important real-world applications such as medical research. Common attacks and information loss possible on anonymized data are intense challenges

[106] The KNN(G,S) clustering algorithm determines the single best neighbour of each cluster and assigns instances one at a time to the existing cluster. The algorithm was modified to overcome the skew in the sensitive value distribution of the resultant clusters using -Nearest Neighbour technique. The KNN-(G,S) clustering algorithm determines the *KN* nearest neighbours using equations from each sensitive value group and adds the *KN* records to the clusters at a time (Vedha Nayahi JJ, 2016). The (G,S) clustering algorithm is proposed to achieve clusters each with S diverse sensitive values for the sensitive attribute. The number of distinct sensitive values in each cluster would be exactly equal to the number of sensitive values in the given data set. The clusters' formed are anonymized by replacing the actual values with the centroid (G) of the cluster.

[107] Lindell Y. and Pinkas, B.,, 2000, "*Privacy Preserving Data Mining*", International Association for Cryptologic Research (IACR), pp 37-55

of anonymization. In this section, the anonymization privacy technique is included here, as a measure for protecting data privacy (not only data storage, as aforementioned, in 5.1.4-IV).

5.1.5.6 Ranking-Oriented Privacy Measures

For privacy apprehensions, secure searches over encrypted cloud data have encouraged numerous research works under the single-owner model. The basics of “Ranking” is precisely explained by Zhang et al. (2016). More specifically, Zhang, in order to efficiently authenticate data users and detect attackers who steal the secret key and perform illegal searches, proposed a novel dynamic secret key generation protocol and a new data user authentication protocol. To enable the cloud server to perform *secure search* among multiple owners’ data encrypted with different secret keys, he systematically constructed a novel secure search protocol. *Ranking* the search results and preserving the privacy of relevance scores between keywords and files, he proposed a novel additive order and privacy preserving function family. Moreover, it showed that this approach is computationally efficient, even for large data and keyword sets.

Additionally, most cloud servers in cutting-edge practice not only serve one owner; instead, they support multiple owners to segment the profits brought by cloud computing. Having multiple owners over the same data is a challenge and it is problematic, especially in sight of the different rights that they might have over this data.

Below, in Table 11 we present our proposed countermeasures for the Data protection countermeasures category, with their best use, strengths and weaknesses. The proposed countermeasures for the Data protection countermeasures category, have been resulted from our research on the literature (see also Ch. 3, Methodology review), which we consider them as the most used.

Table 11: Data protection countermeasures category & their best use cases

Countermeasure	Best For	Strengths	Weaknesses	Value
Cryptography-Oriented (Encryption, Digital Signatures, Homomorphic Encryption)	Confidentiality, Secure Storage, Secure Transactions	Strong security, ensures data secrecy	Computational overhead, key management complexity	*High* (Allows computations on encrypted data without needing to decrypt it, ensuring privacy during processing and mitigating the risk of exposure)
Probability-Oriented (Randomized Response, Noise Injection, K-Anonymity)	Privacy-Preserving Data Analysis	Good for anonymized datasets, protects identity	Risk of re-identification if noise is insufficient	*Moderate* (Introduces noise into datasets used for training models, helping to improve privacy by preventing overfitting and maintaining privacy while still producing useful outputs, though it

				may reduce model accuracy slightly)
Probabilistic Hybrid (Combining Cryptographic & Statistical Techniques)	Secure AI/ML, Encrypted Computing	Balances security and usability	Complex to implement, requires advanced computation	*Moderate* (Combines probabilistic anonymization (e.g., randomization) with cryptographic methods (e.g., cryptography) to protect identity, reducing the likelihood of re-identification while preserving some data utility)
Utility-Preserving (Differential Privacy, Perturbation, Synthetic Data)	Data Analytics, AI Model Training	Keeps data useful while preserving privacy	Hard to balance utility and privacy	*Moderate* (Replaces sensitive data with non-sensitive tokens, allowing for secure data processing while retaining some utility for business operations or analytics, though not all data use cases are supported)
Anonymization (K-Anonymity, L-Diversity, T-Closeness, Generalization, Masking)	Privacy Protection, GDPR Compliance	Reduces risk of direct re-identification	Weak against linkage attacks, privacy loss over time	*High* (Ensures that each data record is indistinguishable from at least **k-1** other records, providing strong anonymization and mitigating the risk of re-identification)
Ranking-Oriented (Data masking, Role-Based Access, Attribute-Based Access)	Authorization	Enforces data access policies	Doesn't protect data if access is breached	*Moderate* (Masks sensitive data but ensures that the relative rankings or order of data remain intact, allowing for analysis while preventing exposure of specific details, making it highly effective in privacy-preserving scenarios).
Basic Password Protection	Small-scale, non-critical systems or legacy environments where	Simple to implement - No additional costs	Vulnerable to brute force and guessing - Susceptible to	*Low* (Basic password protection, such as using simple

	security needs are minimal.	beyond user management	phishing and social engineering	passwords without multi-factor authentication (MFA) or advanced encryption, is a low-effectiveness)
Encryption & Tokenization (E2EE)	Payment Security, Sensitive Data Protection	Strong protection, works well with compliance	Tokenization can be costly; encryption adds processing overhead	*High* (Ensures that data is encrypted on the sender's side and decrypted only on the recipient's side, offering strong privacy protection by preventing unauthorized access at any point in transit or storage.)

5.2 Comparison and results

5.2.1 Key evaluation criteria for cloud privacy countermeasures. A research approach

In section 5.2.1 we will analyze and present the criteria and their role for the evaluation of every category of countermeasures. It is important to note that, for every countermeasure in every category, we present the positive and negative characteristics of its impact. Additionally, we score them with an evaluation value (3-scale grades: “*High*”, “*Moderate*”, “*Low*” based on their general effectiveness) along with justification. The proposed scoring will be used during the selection of the best countermeasure(s), in our proposed 3-phases algorithm (Section 5.2.3.6)

Evaluating countermeasures against cloud computing attacks requires a comprehensive set of criteria that addresses security, efficiency, adaptability, and usability. The right criteria can ensure that countermeasures are not only effective at threat detection and prevention but also practical for implementation in dynamic and complex cloud environments.

Numerous researchers and security organizations have proposed criteria for evaluating countermeasures mainly in cloud privacy, especially in IDS, MiTM, DLPs, Data Storage and Data protection mechanisms (as above). Here are some research institutions and researchers that are influential in defining the best evaluation criteria for cloud privacy countermeasures:

1. **NIST (National Institute of Standards and Technology)**

NIST provides widely referenced guidelines and criteria for cloud security and cloud privacy, including publications like NIST SP 800-53 (Security and Privacy Controls for Information Systems) and NIST SP 500-292 (NIST Cloud Computing Reference Architecture). Evaluation Criteria: NIST’s criteria cover effectiveness, scalability, reliability, and compliance, offering a comprehensive framework for assessing countermeasures. Many researchers use NIST’s framework to evaluate or benchmark cloud security and cloud privacy solutions.

2. ISO/IEC Standards (International Organization for Standardization)

ISO/IEC 27001 and ISO/IEC 27017: These standards provide guidelines for information security management systems (ISMS), security controls specific to cloud services and mainly regarding cloud privacy. Evaluation Focus: ISO/IEC standards emphasize risk assessment, operational performance, resilience, compliance, and usability, making them key resources in cloud security evaluations. Researchers often use these standards to build frameworks that assess cloud countermeasures in various areas, including security, cloud privacy and usability.

3. CSA (Cloud Security Alliance)

Cloud Controls Matrix (CCM): The CSA’s CCM is a popular framework that offers specific security and cloud privacy control recommendations tailored to cloud environments.

STAR (Security, Trust & Assurance Registry): STAR is a registry for assessing the security of cloud providers.

Researchers use CSA’s frameworks to develop criteria that consider effectiveness, compliance, and trustworthiness of cloud countermeasures.

4. Researchers and others (providers, institutions)[108]

Hashizume et al. (2013) outlines critical cloud privacy issues in cloud computing and define essential criteria for cloud privacy countermeasures, including availability, reliability, and scalability. Their work is frequently cited by researchers aiming to address cloud security vulnerabilities through various countermeasures.

Kandukuri et al (2009) focuses on security and compliance issues in cloud computing. They suggest criteria around regulatory compliance, privacy, and reliability, making their work a foundational study for evaluating security controls and countermeasures in cloud environments.

Zhou and Zhang's (2010) survey provides an overview of security and cloud privacy challenges in cloud computing, proposing criteria like data confidentiality, access control, and integrity. Their work has informed many subsequent studies on the effectiveness of cloud security countermeasures.

Jansen and Grance’s (NIST, 2011) publication include criteria for securing public cloud environments and emphasizes usability, scalability, compliance, and availability. It is widely cited in cloud security and privacy research and is a reference for evaluating countermeasures in cloud computing.

Fernandes et. al (2014) provides a comprehensive survey on cloud security challenges and suggest criteria for evaluating countermeasures, including performance, adaptability, and effectiveness against specific attack vectors. This work helps researchers identify important criteria based on current and emerging cloud threats.

[108] See the relevant literature of this section in the final chapter of the thesis: “References”

Almorsy et. al (2016), examine security cloud and privacy challenges specific to cloud environments, emphasizing criteria like adaptability, flexibility, and interoperability for countermeasures. It is valuable for researchers working on flexible, cloud-specific security solutions.

Patel et al (2013), provides a survey of IDS/IPS systems for cloud computing, emphasizing criteria like detection accuracy, real-time response capability, and resource efficiency. Their work is often referenced in the design of cloud-specific intrusion detection and prevention solutions.

ENISA’s report (2009) covers risk assessments and recommendations for cloud privacy, providing criteria for effectiveness, resilience, and adaptability in countermeasures. This document is a reference for European cloud providers and researchers focusing on regulatory and security requirements.

Large cloud service providers (AWS, Google Cloud, Microsoft Azure) as Saas/Cloud providers, often publish best practices and whitepapers detailing security and evaluation criteria including aspects like compliance, resilience, scalability, and cost-effectiveness. These resources are referenced by researchers and practitioners to design or evaluate security countermeasures.

Therefore, in cloud computing security and mainly in cloud privacy research, NIST, ISO/IEC, and CSA are among the most influential organizations, providing comprehensive frameworks that inform the design and evaluation of countermeasures. Researchers like Hashizume, Kandukuri, Zhou, and Zhang (and others cited below) are frequently cited for their studies on cloud security and mainly *cloud privacy* issues and evaluation criteria.

These resources collectively emphasize mainly: *effectiveness*, *performance*, *scalability*, *cost efficiency, integration and compatibility* and *usability,* as the essential criteria for evaluating cloud countermeasures regarding cloud privacy. Researchers and practitioners often build on these foundational works to create specialized frameworks tailored to the unique challenges of securing cloud environments for cloud privacy. In the following, we formally analyze this set of 6 criteria.

1. **Effectiveness**

Effectiveness refers (Yuan X. et al, 2018, Wang, S. et al, 2020) to how well a countermeasure can prevent, detect, or mitigate privacy breaches in the cloud. A countermeasure is considered effective if it can identify unauthorized access, prevent data leakage, or secure communication channels between users and cloud services[109]. It measures how well a countermeasure protects sensitive data, ensures compliance, and mitigates privacy risks in cloud environments. In the effectiveness of a set of countermeasures for cloud privacy and security we have also taken under consideration factors, such as: *Data Protection, Threat Mitigation, Legal and Regulatory Compliance, User and Stakeholder Trust, Resilience to Failures and Attacks.*

[109] (a) Yuan, X., et al. (2018). "*Data Loss Prevention in Cloud Computing: A Review and Classification of Approaches*". International Journal of Computer Science and Information Security.
(b) Wang, S., et al. (2020). "*Encryption in Cloud Security: A Survey*". IEEE Transactions on Cloud Computing.

The *Data Protection* capability refers to the cloud system's ability to secure sensitive data against unauthorized access, loss, or tampering. Effective encryption, access control, and data redundancy mechanisms are part of the data protection capability. Ensuring data protection helps prevent breaches and ensures the confidentiality and integrity of the stored data.

The *Threat Mitigation* capability involves identifying, reducing, and managing potential risks and vulnerabilities in the system. This includes adopting countermeasures like firewalls, intrusion detection systems (IDS), and real-time threat monitoring. The ability to detect, analyze, and mitigate emerging threats ensures that cloud services are resilient against cyberattacks.

The *Legal and Regulatory Compliance* factor ensures that cloud providers comply with data protection laws and regulations like GDPR, HIPAA, and CCPA. Legal and regulatory compliance is essential to avoid legal penalties and maintain the privacy of users' personal and sensitive data within specific jurisdictions.

The *User and Stakeholder Trust* is a key component of cloud privacy. The ability to ensure privacy, confidentiality, and security of data fosters confidence among users, customers, and stakeholders. Transparency in data handling practices and providing users with control over their data can strengthen trust and improve the cloud service provider's reputation.

The *Resilience to Failures and Attacks* refers to the cloud infrastructure's ability to remain functional and maintain privacy and security during attacks or system failures. Resilience is achieved through redundant systems, disaster recovery plans, and robust security protocols, ensuring that privacy is not compromised even during adverse conditions or breach attempts.

2. Performance

Performance refers (Ali, S. et al 2020, Lin F. et al, 2019) to the impact of countermeasures on the overall cloud system’s speed, efficiency, and resource consumption. Cloud privacy countermeasures should not degrade the performance of cloud services, especially in high-demand environments[110]. This ensures that privacy mechanisms not only protect data but also operate efficiently without degrading the system’s usability or scalability. The performance of a set of countermeasures for cloud privacy and security is based mainly on factors, such as : *Latency and Speed (encryption and access time), Resource Utilization (CPU, memory), Throughput, Resilience under Load (stress), Energy Efficiency.*

The *Latency and Speed (Encryption and Access Time)* refers to the time delay (latency) associated with encrypting and decrypting data and accessing stored data. Low latency is crucial for ensuring that privacy-enhancing measures like encryption do not negatively impact the user experience. If encryption slows down access times significantly, it could reduce the efficiency of cloud services. Therefore, balancing strong encryption with fast access is important to maintain both privacy and system performance.

The *Resource Utilization (CPU, Memory)* measures how efficiently a cloud system uses its computational resources, like CPU and memory, when applying privacy measures such as

[110] (a) Ali, S., et al. (2020). "*Performance Evaluation of IDS and DLP Systems in Cloud Security*". International Journal of Cloud Computing and Services Science.
(b) Lin, F., et al. (2019). "*Performance of Encryption Techniques in Cloud Computing: A Benchmarking Approach*". Cloud Computing and Security.

encryption, access controls, and data protection algorithms. Efficient use of resources ensures that privacy solutions do not place an undue burden on the cloud infrastructure, maintaining overall system performance while protecting data.

The *Throughput* refers to the amount of data that can be processed or transferred in a given time period (e.g., megabytes per second). In cloud privacy, throughput is critical for ensuring that privacy measures (such as encryption or data scanning) do not become bottlenecks, allowing the cloud service to scale while maintaining security and compliance.

The *Resilience under Load (Stress)* refers to the ability of a cloud system to handle high traffic volumes, heavy user requests, or stressful conditions while maintaining privacy and performance. This involves ensuring that even during peak usage times or in the face of attacks (e.g., denial-of-service attacks), privacy measures like encryption or access control continue to function effectively without causing system failures or breaches.

The *Energy Efficiency* assesses how well a cloud service can balance privacy protection and resource consumption, minimizing the energy used by servers and data centers. Privacy measures like encryption can be resource-intensive, so optimizing them for energy efficiency ensures that cloud systems remain sustainable, reduce operational costs, and minimize their environmental footprint while maintaining robust data protection.

3. Integration and Compatibility

Integration and *compatibility* refer (Zhou, Y. et al 2019, Mehta, P. et al, 2020) to how easily a countermeasure can be incorporated into existing cloud infrastructures. Countermeasures that are easier to integrate into existing systems are often preferred as they reduce deployment time and costs[111]. Although these criteria are different, they are referenced together in literature, as involved each other. In cloud privacy countermeasures, integration and compatibility are closely related and essential for ensuring that security and privacy solutions work effectively within diverse cloud environments. Both criteria focus on how different privacy-enhancing technologies (PETs), security tools, and compliance frameworks work together to create a seamless, secure, and efficient cloud ecosystem.

Integration determines how seamlessly cloud privacy measures can be incorporated into existing systems, workflows, and cloud environments. A countermeasure with strong integration capabilities ensures smooth implementation, minimizes disruptions, and leverages existing resources effectively. Compatibility ensures that cloud privacy measures work seamlessly with existing systems, applications, and cloud environments. A compatible privacy countermeasure avoids disruptions, reduces implementation complexity, and supports the organization’s broader IT and security strategies.

Regarding the integration and compatibility of a set of countermeasures for cloud privacy and security, we have to take into consideration factors, such as: *Compatibility with Cloud Platforms (Cloud-native support), Integration with Existing Infrastructure, Compliance Integration,*

[111] (a) Zhou, Y., et al. (2019). "*Integrating DLP Solutions with Cloud Security Systems*". Cloud Computing and Security Journal.
(b) Mehta, P., et al. (2020). "*Integration Challenges and Solutions for Cloud Encryption Systems*". Cloud Technology Review.

Integration with Security Ecosystem, Application Compatibility, Operating System and Software Compatibility, Data Format and Protocol Compatibility.

The *Compatibility with Cloud Platforms* (Cloud-native Support), ensures that privacy solutions are optimized for cloud environments such as AWS, Azure, and Google Cloud. At the same time, it supports cloud-native security features like encryption, IAM policies, and privacy controls without requiring extensive modifications.

The *Integration with Existing Infrastructure* ensures privacy solutions can work with on-premises, hybrid, and multi-cloud environments without disrupting operations. It also, show us whether the criterion minimizes configuration challenges and supports legacy systems while maintaining data protection.

The *Compliance Integration* aligns, essentially, with regulatory frameworks such as GDPR, HIPAA, CCPA, and PCI-DSS to ensure legal compliance. It enables automated compliance monitoring and reporting to meet privacy requirements, if succeeded.

The *Integration with Security Ecosystem* works with SIEM (Security Information and Event Management), IAM (Identity and Access Management), firewalls, and threat detection tools. Also, ensures the real-time security monitoring and enforcement of privacy policies across cloud applications.

The *Application Compatibility* supports privacy features across various SaaS, PaaS, and IaaS applications, ensuring that sensitive data remains protected and that enables secure API-based interactions between different cloud services and applications.

The *Operating System and Software Compatibility* ensures that privacy mechanisms function properly across different operating systems (Windows, Linux, macOS) and cloud-based software. Moreover, ensures whether it supports virtual machines (VMs), containers (Docker, Kubernetes), and serverless computing environments.

The *Data Format and Protocol Compatibility* show us how to support diverse file formats (JSON, XML, CSV, etc.) and data transmission protocols (HTTPS, TLS, SFTP, etc.) to ensure secure data handling and to enable smooth data sharing and encryption across platforms without compatibility issues.

4. **Usability**

Usability refers (Patel, R. et. al 2020, Zhang, W. et al, 2021) to how easy it is for administrators to deploy, configure, and manage countermeasures. Usability is particularly important in cloud environments where time to deploy and ease of use directly impact operational efficiency[112]. It determines how effectively users and administrators can interact with and implement these measures without unnecessary complexity or errors. A highly usable privacy countermeasure fosters adoption, reduces operational risks, and improves overall security and privacy outcomes. Regarding the usability of a set of countermeasures for cloud privacy and

[112] (a) Patel, R., et al. (2020). "*Usability Challenges in Implementing DLP Systems in Cloud Environments*". International Journal of Information Security.
(b) Zhang, W., et al. (2021). "*Usability of Encryption Systems in Cloud Security: A Comparative Study*". IEEE Transactions on Cloud Computing.

security, we can take into consideration the following factors, such as: *Ease of Implementation, Intuitive User Interface, Transparency, Minimal Learning Curve (minimal users’ and administrators’ training), Automation and Default Settings.*

The *Ease of Implementation* examines whether privacy solutions should be easy to deploy and configure without requiring extensive modifications to existing infrastructure. Moreover, if it can be supported plug-and-play functionality or pre-built integrations with cloud platforms (AWS, Azure, Google Cloud).

The *Intuitive User Interface (UI)* of a countermeasure should be user-friendly, clear, and easy to navigate, ensuring that users can manage privacy settings without confusion while it should provide dashboards, visual reports, and alerts for better monitoring and control of data privacy.

Under *Transparency*, users should have clear visibility into how their data is being processed, stored, and protected where there is a need to support privacy policies, audit logs, and real-time data access tracking for accountability and compliance.

Considering the *Minimal Learning Curve (Minimal Users’ and Administrators’ Training)*, the system should be easy to learn and use, reducing the need for extensive training for end users and IT teams while it should also provide self-explanatory settings, guided tutorials, and knowledge bases to simplify adoption.

The *Automation and Default Settings* refers to pre-configured default privacy settings which should align with industry’s best practices and compliance standards. Moreover, automation features (e.g., automatic encryption, real-time threat detection, policy enforcement) should reduce manual effort and human error.

5. Cost efficiency

Cost efficiency is a critical consideration (especially for Data Storage protection) when selecting a cloud privacy countermeasure (Zargari, S. et al 2020, Sams, H. et al, 2018). Cloud environments are cost-sensitive, and the benefits of a countermeasure must justify its implementation and maintenance costs[113]. Cost-efficiency assesses the balance between the financial investment required and the level of privacy protection provided. It attempts to ensure that resources are allocated optimally without compromising data security and compliance requirements. A Cost-Benefit Analysis is also necessary, as follows:

- **Value Provided vs. Cost**: Compare the level of privacy protection offered relative to the financial investment. *Example*: Comparing the effectiveness of open-source encryption solutions versus commercial enterprise-grade tools.
- **Return on Investment (ROI)**: Analyze how the countermeasure improves efficiency, reduces risk, or enhances user trust relative to its cost.

  Regarding the cost efficiency criterion for a set of countermeasures for cloud privacy and security is measured, we take into consideration factors, such as: *Initial Implementation*

[113] (a) Zargari, S., et al. (2020). "*Cost-Efficiency in Cloud Security: An Analysis of DLP and IDS*". Journal of Cloud Computing and Security.
(b) Shams, H., et al. (2018). "*Cost-effective Approaches to Implementing Encryption in Cloud Systems*". International Journal of Cloud Computing and Services Science.

Costs, Operational Costs, Compliance and Avoidance of Fines, Cost of Training and Support, Automation and Cost Reduction.

The *Initial Implementation Costs* include expenses for deploying cloud privacy solutions, such as software licensing, infrastructure setup, and integration with existing systems. Also, they can helps in choosing cloud-native, scalable, or open-source privacy tools to assist in reducing upfront costs.

The *Operational Costs* cover ongoing expenses like cloud storage, data encryption, privacy monitoring, and system maintenance. Further, the efficient resource utilization (e.g., optimizing CPU, bandwidth, and storage) by a countermeasure can help minimize long-term costs.

The *Compliance and Avoidance of Fines* investigates the existence of privacy-compliant solutions (e.g., GDPR, HIPAA, CCPA) that help avoid regulatory penalties and lawsuits. Non-compliance can result in hefty fines, reputational damage, and potential business disruptions.

The *Cost of Training and Support* concerns the expenses related to training IT teams, administrators, and employees on privacy best practices and system management. So, when examined, it can unveil user-friendly solutions with minimal learning curves that educe training costs and improve efficiency.

The *Automation and Cost Reduction* factor informs about automated encryption, threat detection, policy enforcement, and compliance reporting which reduce the need for manual intervention. Moreover, if AI-driven privacy solutions are involved, they may lower labor costs by improving efficiency and reducing the risk of human errors.

6. Scalability

Scalability refers (Ghan, H. et al 2019, Sing H, et al, 2021) to the ability of a countermeasure to adapt and perform well as the size and complexity of the cloud environment increases. Scalability is essential for cloud applications as they typically involve many users, resources, and services[114]. It ensures that these measures can adapt to increasing workloads, user bases, and data volumes without compromising performance, effectiveness, or usability. Scalable privacy solutions are critical for businesses operating in dynamic environments or experiencing rapid growth. Regarding the scalability for a set of countermeasures for cloud privacy and security, which is measured, we take under consideration some factors, such as: *Data Volume Handling, User Base Growth, Geographic Expansion, Elasticity (ability to scale-up), Resilience Under Scale, Scalability of Privacy Policies.*

The *Data Volume Handling* ensures that privacy measures (e.g., encryption, access control, and DLP) can manage large and growing datasets efficiently and support big data processing, distributed storage, and cloud-native scaling solutions to maintain performance.

[114] (a) Gan, H., et al. (2019). "*Scalable Intrusion Detection Systems for Cloud Computing*". Journal of Cloud Computing: Advances, Systems, and Applications.
(b) Singh, H., et al. (2021). "*Scalability and Efficiency of DLP Systems in Cloud Security*". Journal of Information Security.

The *User Base Growth* highlights that privacy mechanisms must accommodate an increasing number of users, access requests, and authentication processes and examines whether such mechanisms support identity federation and multi-factor authentication (MFA) for scalable security.

The *Geographic Expansion* ensures that privacy policies and compliance controls remain effective as organizations expand to different regions with varying data protection laws (e.g., GDPR, CCPA, PDPA) and examines whether multi-region data replication, localized compliance enforcement, and cross-border data transfer safeguards are supported.

The *Elasticity (Ability to Scale-Up)* examines whether privacy solutions can automatically scale up or down based on demand, ensuring efficient resource utilization and cost optimization. It also examines whether cloud-native architectures are involved with auto-scaling, containerization (Kubernetes, Docker), and serverless computing support elasticity.

The *Resilience Under Scale* ensures that privacy measures remain effective and performant under high loads, cyberattacks, or system failures and if, under this criterion, the countermeasure’s category uses distributed security mechanisms, failover strategies, and redundancy to prevent data breaches and downtime.

The *Scalability of Privacy Policies* examines whether privacy policies should be easily adjustable as data protection laws, user access levels, and organizational needs evolve. Also, tell us if supports automated policy enforcement, AI-driven compliance monitoring, and adaptive access controls.

Evaluating countermeasures against cloud computing privacy attacks requires, essentially, a comprehensive set of criteria that addresses: *Effectiveness*, *Performance*, *Scalability*, *Cost efficiency, Integration/Compatibility* and *Usability*.

Moreover, evaluating countermeasures for cloud privacy attacks involves considering multiple factors. **Effectiveness** is the most critical factor, followed by **Scalability** and **Cost efficiency**. The **Integration-Compatibility** of solutions are also essential for their successful deployment in diverse cloud environments. **Usability** and **Performance** are often overlooked, but they play a significant role in the day-to-day operations and efficiency of cloud systems. The order could depend on each organisation’s preferences and requirements.

Through research and analysis of the available literature, it is evident that cloud providers are increasingly offering integrated solutions for privacy protection, including IDS, DLP, and encryption. The right combination of these countermeasures, tailored to specific needs and cloud architectures, can provide robust protection against a wide range of privacy attacks.

The right key evaluation criteria can ensure that countermeasures are not only effective at threat detection and prevention but also practical for implementation in dynamic and complex cloud environments. We will use the six criteria (presented in section 5.2.1), namely: ***Effectiveness*****,** ***Performance*****,** ***Scalability*****,** ***Cost efficiency, Integration and Compatibility*** and ***Usability*****,** to compare the main five categories-models (presented in sections 5.1.1-5.1.5) for the five countermeasures’ categories protecting cloud security/privacy: ***IDS, DLPs, Privacy as a Service–Data storage, MiTM*** and ***Data privacy mechanisms.***

5.2.2 Comparative analysis of the main countermeasures categories for cloud privacy attacks based on key evaluation criteria

It is crucial to understand the comparison of suitability of the cloud privacy strategies. The results are presented in Table 12. These results are proposed by the author of the thesis. It is important to realize that the selected key evaluation criteria for the countermeasures' categories, are varying in their "performance" across the mechanisms-categories. Indeed, at the end we assign *tagging* or *evaluation values* as: *"High", "Moderate"* or *"Low",* to every criterion per every countermeasure's category, using a **3 scales scoring**. In some cases, the tagging is characterized by a combination of these values (e.g. "Moderate to High", or "Low to Moderate", etc), when the assessment value for a criterion is specific in multiple cases but might vary in few others towards a different evaluation value. For instance, the performance for a counter-measure category is mainly High but, in some cases, it tends to Moderate. In this case, we assign the more precise value of High to Moderate (e.g., something between "High" and "Moderate", or between "Low" and "Moderate").

The reader can refer to section 5.2.1 (Key evaluation criteria). It is important, also, to make the assumption that the following Table 12 and our relevant analysis (Sections 5.1.1 – 5.1.5 and 5.2.1) concerns mainly the ***Public Cloud.***

5.2.2.1 Comparison countermeasures' categories and relevant tagging

In Table 12, we present a comparison for every countermeasure category for protecting cloud privacy and security from attacks, based on specific key evaluation criteria, tagging them as *"High", "Moderate"* or *"Low".* More specifically:

(I) Effectiveness: How well a countermeasure prevents or mitigates cloud privacy risks.

High Effectiveness: Countermeasures rated as highly effective are capable of consistently addressing the privacy threats they are designed to combat. For instance, IDS is highly effective because it can detect and cover a wide range of intrusions and threats in real-time. Similarly, DLP is highly effective at preventing unauthorized data access or leakage.

Moderate Effectiveness: A moderate effectiveness rating means the countermeasure addresses some, but not all, threats. For example, MiTM protection mechanisms (SSL/TLS encryption, VPNs) are effective at preventing interception during transmission but may not address other threats, such as data leaks within the cloud infrastructure. Data Privacy Protection Mechanisms, like anonymization or pseudonymization, are highly effective for compliance but may not fully secure data against all forms of attack, especially during data processing (with the exception of homomorphic encryption).

Low Effectiveness: A low effectiveness rating means that the countermeasure fails to provide adequate protection, leaving data exposed to risks despite implementation. The countermeasure "Static Encryption Keys" in the MiTM category (if encryption keys are not frequently rotated or managed properly, MiTM attacks can still be successful, leading to low effectiveness in preventing advanced interception techniques) or "Simple Access Control Lists (ACLs)" in the Data Storage category, are such examples.

Justification of the evaluation values of every countermeasure category (Effectiveness criterion)

DLP is highly effective at preventing unauthorized data access and leakage, especially for sensitive information, such as financial records, intellectual property, and personally identifiable information (PII). Advanced DLP solutions use contextual analysis and machine learning to detect and block data exfiltration attempts in real-time (Barkley et al., 2021). Since, the Effectiveness evaluation criterion measures how well a countermeasure protects data from unauthorized access, breaches, and loss while ensuring compliance with privacy regulations, then Data Storage countermeasures can be rated a “high” in effectiveness because they provide strong security guarantees, reliability, and advanced protection mechanisms (**High**)

IDS is highly effective in detecting security breaches, intrusions, and abnormal behaviors within cloud networks (**High**)

MiTM protection mechanisms are just effective in mitigating the interception of sensitive data during transmission (Benson et al., 2020) (**Moderate**)

Data storage security mechanisms are effective at preventing unauthorized access and ensuring data confidentiality (Zhao et al., 2020) (**High**)

Data privacy protection like anonymization, pseudonymization, encryption and differential privacy effectively protect (sensitive) user data while ensuring compliance with data privacy regulations (e.g., GDPR, CCPA) (Janssen et al., 2021). Encryption plays a critical role in data privacy protection, ensuring confidentiality and integrity. Modern encryption methods aim to minimize delays while maintaining strong security. As factors affecting encryption performance for data privacy protection, are considered: Computational Overhead, I/O Latency, Network Throughput[115] (**High**).

(II) Performance: Refers to how efficiently the countermeasure operates in a cloud environment.

High Performance: High-performance solutions have a minimal impact on system functionality and speed. IDS solutions generally perform well in cloud environments, especially cloud-native IDS that are optimized for high throughput and low latency. MiTM protections (e.g., SSL/TLS encryption) also offer high performance because modern encryption algorithms introduce negligible delays.

Moderate Performance: Some countermeasures introduce a trade-off between protection and system efficiency. DLP systems, especially those that analyze content in real-time, can impact performance due to the overhead of scanning large data volumes. Data storage security measures, such as encryption at rest can slightly slow down data access, but advancements in hardware and cloud solutions have reduced this impact.

[115] For type AES-256 Security is High and Speed is Fast, type RSA-4096 Security is High and Speed is Slow, type ECC Security is somehow High and Speed is Faster than RSA, type ChaCha20 Security is somehow High and Speed is Very Fast and type AES-256 Security is High and Speed is Very Fast.

Low Performance: Low performance operates with remarkable impact on system functionality and speed, so significantly degrades system efficiency, leading to slow processing, high latency, or increased resource consumption. The "TLS Inspection for Network DLP" countermeasure in the DLP category (decrypting and inspecting encrypted traffic introduces processing delays and requires substantial computational resources) and "Homomorphic Encryption" in Data Storage category (computationally expensive), are such examples.

Justification of the evaluation values of every countermeasure category (Performance criterion)

Performance may be impacted by the data inspection overhead, especially in high-volume environments. However, cloud-based *DLP solutions* optimize performance using distributed resources to minimize latency (Bhuyan et al., 2020) (**Moderate**)

Cloud-based *IDS systems* are optimized to handle high traffic loads with minimal performance degradation, leveraging the cloud's distributed resources. However, performance can decrease if the IDS is not properly tuned (Venkatesh et al., 2020)[116]. HIDS sometimes introduces delays (**Moderate**).

Encryption mechanisms, in *MiTM*, can introduce slight delays in communication, but the performance impact is characterized with some delays (although sometimes minimal) with modern cryptographic techniques and optimized cloud solutions (**High to Moderate**).

Data encryption can impact *data storage* performance, especially when dealing with large datasets, but modern encryption techniques (hardware-accelerated encryption (AES-NI), optimized cryptographic libraries, and transparent encryption mechanisms) minimize the delay (Zhou et al., 2021) (**Moderate**).

Privacy-enhancing techniques like anonymization and pseudonymization may have a impact on performance due to the need for additional processing, in *data privacy protection* mechanisms. However, cloud platforms often optimize these techniques for minimal performance loss (Yang et al., 2020) (**Moderate**).

(III) Scalability: The ability of a countermeasure to handle growth in data, users, or traffic volume without degradation of performance.

High Scalability: Since Scalability is a key factor in cloud privacy countermeasures, ensuring that security and privacy controls adapt seamlessly to growing workloads, user demands, and evolving threats. High Scalability requires full adaptation of user demands and handle growth in data Cloud-native solutions, such as IDS and MiTM protection mechanisms, are designed to scale seamlessly with the cloud. These solutions can adjust to growing traffic volumes and data sizes dynamically. Data storage security mechanisms like encryption at scale are easily implemented in distributed storage systems, offering elasticity.

[116] It might also rely on the type of IDS employed (e.g., NIDS vs HIDS).

Moderate Scalability: Some countermeasures face few challenges in scaling effectively, especially on cloud environments (adapting also to legacy environments where necessary). DLP, for example, might struggle with large-scale deployments because real-time data inspection requires significant resources. While cloud DLP solutions are scalable, they may still require tuning for performance when data volumes increase.

Low Scalability: A low scalability rating means that the countermeasure faces multiple, strong challenges in scaling effectively. In particular, it struggles to adapt to growing cloud demands, leading to bottlenecks, inefficiencies, and potential security risks. The “Manual Rule-Based Content Inspection” countermeasure in the DLP category (manually configured DLP rules become difficult to manage and scale) and “Secure Multi-Party Computation” in the Data privacy prevention category (requires heavy computation and communication between multiple parties) are such examples.

Justification of the evaluation values of every countermeasure category (Scalability criterion)

DLP solutions are scalable, especially in cloud-based systems, which can dynamically scale to accommodate growing data storage and network traffic (**High to Moderate**)

IDS solutions are highly scalable in the cloud. They adapt to growing cloud infrastructures by leveraging cloud elasticity to handle increasing network traffic and data storage (Patel et al., 2019) (**High**).

MiTM protection mechanisms are scalable in cloud environments, with cloud services like VPNs offering auto-scaling capabilities to match network demands (Li et al., 2020) (**High to Moderate**).

Data storage security solutions are highly scalable, especially in cloud environments where they can expand alongside growing data needs. Encryption and access control policies can be applied across large-scale distributed storage systems (Kamble et al., 2020) (**High**).

Data privacy protection mechanisms scale well in cloud environments, especially when integrated with cloud-native data protection services that automatically apply privacy controls across distributed datasets (Venkatesh et al., 2020) (**High**).

(IV) Cost Efficiency: How affordable the countermeasure is in terms of initial setup and ongoing maintenance. This case plays an important role in the Data Storage protection countermeasure category.

High-Cost Efficiency: Cloud privacy solutions must balance security, scalability, and cost-efficiency to be sustainable. High-Cost efficient countermeasures enhance cloud privacy while minimizing expenses. Solutions, such as IDS and MiTM protection mechanisms (VPNs, TLS), can be cost-effective, especially when offered as cloud-native services with pay-per-use pricing. Data Privacy Protection Mechanisms, especially anonymization or pseudonymization, can also be cost-efficient when implemented in compliance with privacy regulations.

Moderate Cost Efficiency: DLP and Data Storage Security mechanisms, particularly those involving encryption and key management, can incur higher costs (than IDSs). Encryption at rest requires processing power, and DLP systems require resources for scanning large amounts of data, making them potentially more expensive as data volumes grow.

Low-Cost Efficiency: A countermeasure with low-cost efficiency means that it requires high investment but does not provide proportional security benefits, leading to budget constraints, unnecessary expenses, or financial inefficiencies. Thus, it leads to very high costs. The countermeasure “AI/ML-Based IDS” in the IDS category (while effective, these systems require high computational resources, frequent model training, and specialized personnel, leading to high costs) and the “End-to-End Encryption with Frequent Key Rotation” countermeasure in the MiTM category, are such examples.

Justification of the evaluation values of every countermeasure category (Cost Efficiency criterion)

Cloud-based *DLP systems* are cost-efficient, offering flexible pricing models, such as pay-as-you-go, but they can become expensive as data volume and complexity grow (Shams et al., 2020) (**Moderate**).

IDS solutions often have a pay-as-you-go pricing model in cloud environments, providing cost efficiency. However, high-end threat detection capabilities can add to operational costs (Shams et al., 2021) (**Moderate to High**).

While effective, encryption-based *MiTM* protection mechanisms can be resource-intensive and may incur additional costs for key management and encryption services (Mishra et al., 2021) (**Moderate to High**).

Costs are associated with encrypting large volumes of data, but cloud services typically provide cost-efficient models based on *data storage* consumption. However, encryption and key management may add to operational costs and vary significantly (Chen et al., 2021). High costs (Low ‘Cost Efficiency’) are justified due to computational complexity, resource consumption, while moderate costs are justified for some countermeasures, because compliance requirements offer a balance between security and affordability. Low costs (High ‘Cost Efficiency’) countermeasures in cloud data storage protection which often require high initial investment in security infrastructure, automation, and policy enforcement. While basic encryption, access controls, and monitoring tools minimize ongoing costs, they may lack advanced threat protection. Organizations must weigh security needs against budget constraints, opting for a cost-effective mix of encryption, tokenization (**Low, Moderate** or **High**).

Cloud data privacy protection mechanisms tend to incur moderate costs due to the need for additional processing and specialized tools. Cloud-native solutions can mitigate some of these costs by providing integrated, scalable options (Mehta et al., 2021). Such

options which are integrated make the tools scalable, and with automated security features[117] (**Moderate**).

(V) Integration and Compatibility: How well a countermeasure integrates with other security systems and cloud infrastructure.

High Integration and Compatibility: Countermeasures (categories) are *High* in Integration/Compatibility when they have seamless compatibility with cloud platforms and suitable interoperability with existing infrastructure. IDS integrates easily with other security layers, such as firewalls and SIEM systems, while MiTM protection integrates well with cloud security protocols, SSL/TLS, and VPNs. Data Privacy Protection Mechanisms are also designed to integrate well with modern cloud environments, supporting regulatory compliance and data protection standards.

Moderate Integration and Compatibility: A countermeasure with moderate integration/compatibility may lack support for integration with one or more cloud platforms, third-party tools or applications or might have one or more incompatibilities with existing platforms or tools that prevent its deployment. DLP solutions, while integrating well with cloud platforms and endpoints, may face challenges when integrated with some third-party applications. Data Storage Security mechanisms like encryption might face compatibility issues with non-cloud-native storage systems or older applications that do not support modern encryption standards. Examples of moderate integration & compatibility countermeasures, meaning they require some adjustments, API connections, or policy modifications for full deployment are: End-to-End Encryption (E2EE) with BYOK & HYOK (Hold Your Own Key). These are examples of countermeasures that require integration with cloud provider encryption tools.

Low Integration and Compatibility: A countermeasure with low integration/compatibility is difficult to deploy, lacks support for multiple cloud platforms, or requires extensive modifications to function properly. The countermeasure "Strict Network Segmentation with Legacy Firewalls" in MiTM category (Traditional firewalls may struggle to integrate with cloud environments due to dynamic IP allocation and encrypted traffic) and the "Hardware-Based Encryption Modules" countermeasure in the Data Storage category, are such example.

Justification of the evaluation values of every countermeasure category (Integration and Compatibility criterion)

DLP solutions integrate well, in general, with cloud services, security information and event management (SIEM) systems, plus endpoint protection tools, as authors claim (Yang et al., 2019). Based on current research, modern DLP solutions are highly integrable and cloud-compatible[118] (**High to Moderate**).

[117] AWS KMS, Azure Key Vault, etc

[118] However, some challenges and limitations of DLP in the Cloud, might be, as : limited visibility in encrypted or BYOK (Bring Your Own Key) scenarios, lack of full multi-cloud consistency or performance and latency concerns.

Integration with other security mechanisms like firewalls, encryption tools, and cloud service platforms is seamless, particularly for cloud-native *IDS systems* (Zhang et al., 2020). Regarding the final evaluation values, we consider optimal scenarios, with the sense of having cloud-native mechanisms (**High**).

Most *MiTM* protection mechanisms are highly compatible with cloud environments, especially those that support encryption standards like SSL/TLS. However, older systems may require upgrades to fully integrate with modern cloud encryption protocols (Zhang et al., 2019)[119] (**High**).

Data storage security mechanisms for cloud, on-premises, and hybrid storage environments, need top security solutions (meaning such solutions as they provide features that can be well integrated with the countermeasures) in order to ensure strong integration and compatibility across cloud platforms (such as AWS, Google Cloud, and Microsoft Azure). Such software solutions can provide built-in security features like encrypted storage volumes (Li et al., 2020)[120] (**High to Moderate**).

Cloud data privacy protection tools may integrate adequately with cloud environments that support modern privacy regulations and compliance tools. These tools integrate specific privacy protection tools into cloud environments (like AWS, Azure, and Google Cloud) for data privacy, but some older cloud platforms may require additional configuration to meet privacy requirements (Benson et al., 2020) (**High**).

(VI) Usability: How easy it is for users (including administrators) to deploy, configure, and use the countermeasure. Essentially, assess how easy, efficient, and intuitive a security solution is for users and administrators to implement, manage, and operate. It also assesses whether a countermeasure can be effectively utilized without extensive technical expertise or complex configurations.

High Usability: We characterize a countermeasure as “High" in the Usability criterion, if it is easy to implement, intuitive for users and administrators, requires minimal training, and automates key security functions. Ease of implementation (as “Cloud-native security tools”) and Minimal Learning Curve (as “Automated encryption tools”) characterize the countermeasures in such case. MiTM protection mechanisms, such as SSL/TLS and VPNs are relatively simple to configure and deploy, particularly with cloud service providers who offer integrated security tools. IDS systems, especially cloud-native ones, come with user-friendly interfaces and automated threat detection that require minimal intervention.

Moderate Usability: A countermeasure has “moderate” usability when there are some small difficulties to implement it and/or might require some basic, potentially periodic training in order to be properly used. DLP solutions require configuration and ongoing monitoring, which might be more complex, especially in large organizations. Data storage

[119] VPNs (Virtual Private Networks) are often integrated well to protect against MiTM attacks, but they are not always foolproof. In some cases, attackers can still perform MiTM attacks against or even through a VPN connection.

[120] For on-premise storage security solutions, might use BitLocker Windows, Veritas Data Insight, etc.

security mechanisms like encryption require proper management of encryption keys, which can add to the complexity, making them moderately usable.

Low Usability: A countermeasure with low usability is difficult to configure, complex to operate, or requires extensive expertise, making it impractical for real-world use. The countermeasure “Mutual TLS (mTLS) for End-to-End Encryption” in the MiTM category (complex certificate management increases administrative overhead) and the “Secure Multi-Party Computation” in the Data Privacy prevention category, are such examples.

Justification of the evaluation values of every countermeasure category (Usability criterion)

DLP solutions are relatively user-friendly but may require training for administrators to configure policies effectively and manage alerts (**Low to Moderate**).

Cloud-native *IDS systems* are relatively easy to use, with user-friendly interfaces and automated threat detection. Traditional IDS systems may require more manual configurations and expertise to be properly managed (Mehta et al., 2021). We can define rules for the usability of the IDS as: how easily security teams can deploy, configure, interpret alerts, and take action (namely, Ease of Deployment and Configuration, Intuitive User Interface (UI) and Dashboard, Response Automation and Actionability) (**High**).

Implementing *MiTM* protection mechanisms in the cloud is generally straightforward for cloud providers and users, though proper key management and certificate handling can be complex (**High**).

Data storage security mechanisms are relatively easy to configure but may require careful management of access control policies and encryption keys. Cloud providers offer user-friendly interfaces for managing these settings (**Moderate**).

While the implementation of *data protection privacy* mechanisms can be complex, modern cloud platforms simplify the process with user-friendly compliance dashboards and automated privacy controls. Automated privacy controls all cases but with a complex configuration (**Moderate**).

We present our proposed Table 12, crossing the six evaluation criteria with the five countermeasure categories, giving the respective evaluation value (tagging), as analyzed in sections 5.1.1 – 5.1.5.

Table 12. Comparative table: Evaluation criteria for cloud privacy countermeasures (concerns mainly the public cloud)

Countermeasures' categories for cloud privacy protection from attacks					
Criteria	**DLPs**	**IDS**	**MiTM Protection**	**Data Storage Security**	**Data Privacy Protection**
Effectiveness	*High*: Prevents unauthorized data access and all other attacks - threats	*High*: Detects intrusions and anomalies	*Moderate*: Protects against data interception	*High*: Prevents unauthorized access to data and all other attacks - threats	*High*: Ensures privacy compliance
Performance	*High to Moderate*: Data inspection overhead	*High to Moderate*: May cause slight latency (HIDS)	*High to Moderate*: Minimal delay with modern encryption	*Moderate*: Performance impact with encryption	*Moderate*: Processing overhead
Scalability	*High to Moderate*: Easily scalable with cloud growth but may require tuning	*High*: Cloud-native IDS scale dynamically	*High to Moderate*: Scalable with cloud VPNs	*High*: Scalable encryption and access control	*High*: Scalable across cloud data
Cost Efficiency	*Moderate*: Pay-as-you-go but can be costly with large data	*Moderate to High*: Cloud-native pricing models	*Moderate to High*: Costs for encryption and key management	*Low, Moderate or High*: Storage and encryption costs	*Moderate*: Additional processing and tools (for data masking, encryption storage, access control, secure data sharing)
Integration & Compatibility	*High to Moderate*: Integrates with cloud and endpoint systems. May face some challenges (data fragmentation, latency issues, diverse cloud architectures)	*High*: Seamless with cloud services	*High*: Integrates with SSL/TLS and cloud encryption	*High to Moderate*: Integrates with cloud platforms, but with some compatibility issues (data format, schema & data model mismatches, API & protocol)	*High*: Easily integrates with cloud compliance tools and the other tools for the attacks - threats
Usability	*Low to Moderate*: Complicated policies that require fine-tuning for different data types. High false positive rates disrupt workflows.	*High*: relatively easy to use, with user-friendly interfaces and automated threat detection.	*High*: Simple once configured (not so frequent reconfiguration)	*Moderate*: Easy for cloud environments but requires encryption key management	*Moderate*: Automated privacy controls for all cases but complex configuration

5.2.2.2 Conclusion

Each cloud privacy countermeasure (DLPs, IDS, MiTM protection, Data Storage Security, and Data Privacy Protection) plays a critical role in ensuring cloud data security and privacy. While cloud-native solutions excel in scalability, integration, and effectiveness, they may still face

challenges related to performance, cost, and usability. Understanding these criteria is crucial for selecting the right countermeasures based on the specific needs of a cloud environment.

This analysis highlights the strengths and weaknesses of various cloud privacy countermeasures (DLPs, IDS, MiTM protection, Data Storage Security, and Data Privacy Protection) based on common evaluation criteria. While protection mechanisms tend to excel in usability, scalability, and integration, IDS and MiTM excel mainly in usability and integration. The choice of countermeasure ultimately depends on data volume, integration requirements plus the relative importance between effectiveness, performance and usability. But in any case, it seems that we need to combine countermeasures in many categories to achieve high values in many criteria.

Data storage and Data protection have classified Moderate performance in Usability while DLP Low to Moderate such that we can consider that the support for this criterion should be improved. However, the support to cost-efficiency needs even further improvement, as the performance of the countermeasure categories in this criterion is mainly moderate.

5.2.3 Towards a countermeasure selection approach

Below, we analyze the suitability of different cloud environments (Public, Private, Hybrid[121], and Multi-Cloud[122]) for all countermeasures’ categories based on the six evaluation criteria: *effectiveness, performance, scalability, cost efficiency, integration and compatibility,* and *usability,* tagging them as *high* or *moderate* or *moderate to high,* or *high to moderate* or *low (*or even as *limited)*. The suitability is translated into differentiated performance values of each countermeasure category across the different criteria and cloud environment types.

In our approach we will focus on the *Multi-Cloud environment* (with Public, Private, Hybrid environments) instead of the beforementioned *Community Cloud*, as multi-cloud environment provides flexibility, resilience, and cost savings although it comes with management complexities. Compared to cloud community environments, multi-cloud strategies offer greater independence, enhanced security, and better performance. Organizations must adopt best practices, leverage automation, and implement strong security controls to fully benefit from multi-cloud strategies. By doing so, they can achieve greater business agility, improved security posture, and optimal cloud resource utilization.

5.2.3.1 Suitable Cloud Environments for DLP countermeasures

Data Loss Prevention (DLP) countermeasures are vital for maintaining data security and preventing unauthorized access, particularly in cloud environments. The effectiveness of DLPs depends significantly on the cloud environment in which they are deployed.

[121] *Hybrid clouds* work by combining the resources and services from two or more separate computing environments. Hybrid cloud architectures require integration, orchestration, and coordination so you can share, shift, and synchronize information quickly

[122] *Multi-cloud* is when an organization uses cloud computing services from at least two cloud providers to run their applications. Instead of using a single-cloud stack, multi-cloud environments typically include a combination of two or more public clouds, two or more private clouds, or some combination of both

Research by Peter Mell and Tim Grance[123] (NIST, 2011) has had significant influence in shaping security and cloud privacy evaluation criteria for cloud computing. Their important contribution is the deployment models of the different cloud environments (Public, Private, Hybrid, Multi-Cloud).It defines essential cloud characteristics, such as resource pooling, elasticity, and service models, that affect how countermeasures are assessed.

Below is a detailed analysis based on the six evaluation criteria: *effectiveness, performance, scalability, cost efficiency, integration and compatibility,* and *usability* (Table 13).

Table 13: DLPs category countermeasures for cloud environments based on key evaluation criteria

	Cloud environments			
Criteria	**Public Cloud**	**Private Cloud**	**Hybrid Cloud**	**Multi-Cloud**
Effectiveness	High	High	High	Moderate to High
Performance	High to Moderate	High	Moderate	Variable
Scalability	High to Moderate	Limited	Moderate	High
Cost Efficiency	Moderate	Moderate	Moderate	Variable
Integration & Compatibility	High to Moderate	High	Moderate	Variable
Usability	Low to Moderate	High	Moderate	High
Best use cases	Dynamic and scalable needs	Regulatory compliance	Mixed sensitivity data	Flexibility and redundancy

(I) Public Cloud

Public cloud environments are shared infrastructures hosted by third-party providers, such as AWS, Google Cloud, or Microsoft Azure.

Suitability for DLP:

- Strengths:
  - Native integration with DLP solutions offered by cloud providers (also a DLP solution with another one), such as AWS Macie[124], Azure Information Protection, or Google Cloud DLP.
  - High (in general) scalability, efficiency and performance to handle large-scale data processing.
  - Cost efficiency through pay-as-you-go pricing models.
- Challenges:

[123] Mell, P. and Grance, T. (2011), "*The NIST Definition of Cloud Computing*", Special Publication (NIST SP), National Institute of Standards and Technology, Gaithersburg, MD, [online], https://doi.org/10.6028/NIST.SP.800-145, p.3, Computer Security Division Information Technology Laboratory, National Institute of Standards and Technology, US Department of Commerce, NIST Special Publication 800-145

[124] https://aws.amazon.com/macie/

- Shared infrastructure increases the risk of data exposure due to multi-tenancy-related threats.
- Somehow limited control over the physical infrastructure and usability.

Best Use Cases:

- Organizations seeking cost-effective, scalable DLP solutions without requiring direct control over the infrastructure.
- Startups or businesses with dynamic workloads and limited IT budgets.

(II) Private Cloud

Dedicated infrastructure is hosted either on-premises or by a third party exclusively for a single organization.

Suitability for DLP:

- Strengths:
  - High effectiveness due to tighter control over infrastructure and data management policies.
  - Superior compliance capabilities for industries require stringent regulatory adherence, such as healthcare and finance.
- Challenges:
  - Limited scalability compared to public clouds.
  - Higher costs associated with infrastructure maintenance and as well as DLP deployment.

Best Use Cases:

- Enterprises handle highly sensitive data, such as intellectual property or confidential customer records.
- Organizations requiring full control over data flow and security configurations.

(III) Hybrid Cloud

A combination of public and private cloud environments with data and applications shared between them.

Suitability for DLP:

- Strengths:
  - Flexibility to store sensitive data in private clouds while leveraging the scalability of public clouds for less critical workloads.
  - Supports tailored DLP solutions that align with specific data sensitivity levels.
- Challenges:
  - Complexity in managing and integrating DLP policies across different environments.
  - Potential latency issues during data transfer between environments.

Best Use Cases:

- Organizations require both high security and cost-efficient scalability.
- Businesses operating in diverse regulatory environments with varying compliance requirements.

(IV) Multi-Cloud

Utilization of multiple cloud service providers for different applications or functions.

Suitability for DLP:

- Strengths:
  - Enhanced flexibility and redundancy by spreading data and workloads across multiple platforms.
  - Ability to leverage specialized DLP features from various providers.
- Challenges:
  - Complex management and integration of DLP policies across multiple platforms.
  - Increased costs and potential integration issues between providers.
  - Latency issues due to the transfer of data between clouds as is the case in hybrid cloud

Best Use Cases:

- Large enterprises with global operations and diverse cloud needs.
- Organizations aim to reduce vendors lock-in and improve resilience.

The ideal cloud environment for *DLP countermeasures* depends on the organization's specific needs:

- **Public Cloud**: Suitable for scalability and cost-effectiveness but less control-intensive scenarios (performance).
- **Private Cloud**: Best for organizations prioritizing data control and regulatory compliance.
- **Hybrid Cloud**: Ideal for businesses requiring a balance between security and scalability.
- **Multi-Cloud**: Best for flexibility and avoiding vendor lock-in but complex to manage.

Organizations must weigh their priorities—whether cost, scalability, control, or compatibility—to determine the most suitable environment for deploying DLP countermeasures.

5.2.3.2 Suitable Cloud Environments for IDS countermeasures

Intrusion Detection Systems (IDS) are vital for identifying unauthorized access and monitoring suspicious activities within cloud environments. The suitability of IDS countermeasures varies across different cloud models—Public, Private, Hybrid, and Multi-Cloud—based on their unique characteristics, operational requirements, and deployment scenarios. Below is a detailed analysis based on the six evaluation criteria: *effectiveness, performance, scalability, cost efficiency, integration and compatibility,* and *usability* (Table 14).

Table 14: IDS category countermeasures for cloud environments based on key evaluation criteria

	Cloud environments			
Criteria	**Public Cloud**	**Private Cloud**	**Hybrid Cloud**	**Multi-Cloud**
Effectiveness	High	High	High	Moderate to High
Performance	High to Moderate	High	Moderate to High	Moderate
Scalability	High	Limited	Moderate to High	High
Cost Efficiency	Moderate to High	Low to Moderate	Moderate	High
Integration & Compatibility	High	High	Moderate	Moderate
Usability	High	Moderate to High	Moderate	Low to Moderate
Best Use Cases	General-purpose use cases	High-security internal operations	Mixed sensitivity workloads	Distributed, redundant operations

(I) Public Cloud

Public clouds are operated by third-party providers and share infrastructure among multiple tenants. Examples include AWS, Google Cloud, and Microsoft Azure.

Suitability for IDS Countermeasures:

- Strengths:
  - Public cloud providers often offer built-in IDS solutions as part of their cloud security suite, such as AWS GuardDuty[125] or Azure Security Center.
  - High scalability allows efficient monitoring of fluctuating workloads.
  - Regular updates and integration with cloud-native tools improve the effectiveness of IDS deployment.
  - Also its high usability, effectiveness and compatibility, affects its wider acceptance.
- Challenges:
  - Shared infrastructure can lead to noisy alerts and difficulty distinguishing between legitimate and malicious traffic and probably affects some the performance. Some cost efficiency issues in order to achieve higher performance due to enhanced and required infrastructure.
  - Limited visibility into the underlying hardware and network configurations may restrict custom IDS tuning.

Best Use Cases:

- Organizations seeking out-of-the-box IDS solutions with minimal setup.
- Suitable for e-commerce platforms, SaaS applications, and small to medium enterprises.

(II) Private Cloud

[125] https://aws.amazon.com/guardduty/

A private cloud is a dedicated infrastructure designed exclusively for a single organization.

Suitability for IDS Countermeasures:

- Strengths:
    - Full control over infrastructure allows fine-tuning and customization of IDS systems.
    - Lower risk of noisy alerts due to dedicated resources and controlled environments.
    - Enhanced ability to integrate IDS with other on-premises security measures.
- Challenges:
    - Higher deployment and maintenance costs due to specialized hardware and software.
    - Limited scalability compared to public cloud models.

Best Use Cases:

- Industries with stringent compliance and security needs, such as finance, healthcare, and government.
- Suitable for organizations prioritizing data sovereignty and high-sensitivity workloads.

(II) Hybrid Cloud

Combines public and private cloud infrastructures, allowing data and applications to move between them.

Suitability for IDS Countermeasures:

- Strengths:
    - Flexibility to monitor traffic across both public and private segments ensures comprehensive threat detection.
    - IDS can be deployed in the private segment to analyze sensitive data while leveraging public cloud scalability for monitoring general traffic.
    - Enhanced adaptability to varying workload security requirements.
- Challenges:
    - Increased complexity in synchronizing IDS deployments across multiple infrastructures.
    - Latency and configuration issues may arise during data movement between public and private environments.

Best Use Cases:

- Enterprises manage both sensitive and non-sensitive data workloads.
- Suitable for organizations needing flexible and scalable IDS solutions, such as multinational corporations.

(IV) Multi-Cloud

Involves using multiple cloud providers for various services, enabling redundancy and avoiding vendor lock-in.

Suitability for IDS Countermeasures:

- Strengths:
  - Redundancy and diversification reduce the risk of a single point of failure.
  - Ability to deploy IDS solutions tailored to the specific capabilities of each provider.
  - Greater resilience in monitoring distributed or global operations.
- Challenges:
  - Managing multiple IDS deployments across providers increases complexity.
  - Inconsistent standards between cloud providers can complicate IDS configuration.

Best Use Cases:

- Organizations with geographically distributed operations or diverse workloads.
- Suitable for businesses prioritizing high availability and risk diversification.

The choice of cloud environment for IDS countermeasures depends on the organization's specific requirements:

1. **Public Cloud**: Best for general-purpose use cases, small to medium businesses, and organizations looking for built-in IDS solutions with minimal setup.
2. **Private Cloud**: Suitable for enterprises requiring high customization and control, especially in compliance-heavy industries.
3. **Hybrid Cloud**: Ideal for businesses balancing sensitive and less critical workloads.
4. **Multi-Cloud**: Optimal for global enterprises requiring flexibility and redundancy.

The decision should align with the organization's scalability needs, security requirements, and budget constraints, ensuring effective monitoring and threat mitigation across the cloud ecosystem.

5.2.3.3 Suitable Cloud Environments for *MiTM countermeasures*

Man-in-the-Middle (MiTM) attacks are a significant threat to cloud environments, exploiting vulnerabilities in communication channels to intercept or manipulate data. Effective countermeasures depend on the cloud environment's structure and operational requirements. Below, we analyze the suitability of various cloud environments—Public, Private, Hybrid, and Multi-Cloud—for deploying MiTM countermeasures (based on the six evaluation criteria: *effectiveness, performance, scalability, cost efficiency, integration and compatibility,* and *usability* (Table 15).

Table 15: MiTM category countermeasures for cloud environments based on key evaluation criteria

	Cloud Environments			
Criteria	**Public Cloud**	**Private Cloud**	**Hybrid Cloud**	**Multi-Cloud**
Effectiveness	Moderate	High	High	Moderate to High
Performance	High to Moderate	High	Moderate to High	Moderate
Scalability	High to Moderate	Limited	Moderate	High

Cost Efficiency	Moderate to High	Low to Moderate	Moderate	Moderate
Integration & Compatibility	High	High	Moderate	Low to Moderate
Usability	High	Moderate	Moderate	Low to Moderate
Best Use Cases	General-purpose use cases	Highly sensitive data	Mixed sensitivity data	Distributed workloads

(I) Public Cloud

Public cloud environments are operated by third-party providers and host multiple tenants on shared infrastructure.

Suitability for MiTM Countermeasures:

- Strengths:
  - Cloud providers often offer built-in security measures, such as SSL/TLS encryption, network segmentation, and MFA to prevent MiTM attacks.
  - Scalability and frequent updates (with high compatibility) to security protocols improve protection against evolving threats.
- Challenges:
  - Shared infrastructure increases exposure to network vulnerabilities.
  - Limited direct control over infrastructure may hinder custom countermeasure implementation that may lead to moderate performance.
  - Moderate in cost-effectiveness in multi-cloud environments (require strong security controls)

Best Use cases:

- End-to-end encryption (e.g., SSL/TLS) for all communications.
- Multi-factor authentication to secure session initiation.
- DNS Security Extensions (DNSSEC) to prevent DNS spoofing (but implementation varies by provider: AWS Route 53[126], Google Cloud DNS, Cloudfare[127]).

(II) Private Cloud

A private cloud is dedicated infrastructure for a single organization, either hosted on-premises or by a third party.

Suitability for MiTM Countermeasures:

- Strengths:
  - High control over security configurations allows for robust and tailored MiTM countermeasures.

[126] https://aws.amazon.com/route53/

[127] https://www.cloudflare.com/learning/dns/dns-cache-poisoning/

  - Reduced exposure to external threats compared to public clouds.
- Challenges:
  - Higher costs associated with deploying advanced security solutions.
  - Limited scalability compared to public clouds.

Best Use cases:

- Implementation of robust encryption standards for internal communications.
- Deployment of firewalls, MFA and secure DNS.
- Use of private key infrastructure (PKI) for secure authentication mechanisms.

(III) Hybrid Cloud

A hybrid cloud combines public and private clouds, allowing data and applications to move between them.

Suitability for MiTM Countermeasures:

- Strengths:
  - Flexibility to apply MiTM countermeasures to sensitive data in the private segment while using public cloud scalability for non-sensitive workloads.
  - Ability to establish secure communication channels between public and private clouds.
- Challenges:
  - Increased complexity in managing and synchronizing security measures across environments.
  - Risk of misconfigurations during data transfers, and also compatibility issues between the different countermeasures in the public and private clouds used.

Best Use cases:

- Implementation of Virtual Private Networks (VPNs) for secure connections between private and public segments.
- Secure API gateways for controlled data exchanges.
- Use of strong encryption protocols and session timeouts to protect cross-cloud communications. Also, MFA and secure DNS should be employed.

(IV) Multi-Cloud

Multi-cloud involves the use of multiple cloud providers for various services or applications.

Suitability for MiTM Countermeasures:

- Strengths:
  - Redundancy and distributed architecture reduce the impact of a single compromised communication channel.
  - Ability to utilize specialized MiTM countermeasures offered by individual providers.

- Challenges:
    - Complexity in ensuring uniform security measures across different platforms.
    - Potential vulnerabilities due to inconsistent encryption and authentication standards (applied in the different clouds exploited).

Recommended MiTM Countermeasures in Multi-Cloud:

- Centralized security orchestration to enforce consistent encryption and access control policies.
- Mutual TLS authentication between services.

The choice of cloud environment for deploying *MiTM countermeasures* depends on the organization’s risk profile, scalability needs, and budget constraints:

- **Public Cloud**: Suitable for organizations relying on cloud-native security features and scalability but requiring minimal custom security measures.
- **Private Cloud**: Ideal for enterprises prioritizing data control and stringent MiTM protections, such as in healthcare or finance.
- **Hybrid Cloud**: A balanced solution for businesses managing both sensitive and less-critical data.
- **Multi-Cloud**: Best for organizations aiming to mitigate vendor lock-in while ensuring redundancy and flexibility but prepared to manage complex security configurations.

Organizations must tailor their MiTM countermeasures to their chosen environment to mitigate risks effectively and ensure secure communication.

5.2.3.4 Suitable Cloud Environments for Data Storage countermeasures

Cloud environments play a critical role in securing data storage, with each type offering unique characteristics that affect the effectiveness of countermeasures. Data storage countermeasures, such as encryption, access controls, redundancy, and backup strategies, need to align with the strengths and limitations of different cloud models: Public, Private, Hybrid, and Multi-Cloud. Below is an analysis of these environments for data storage countermeasures (Table 16).

Table 16: Data storage (Privacy as a Service) category countermeasures for cloud environments based on key evaluation criteria

	Cloud environments			
Criteria	**Public Cloud**	**Private Cloud**	**Hybrid Cloud**	**Multi-Cloud**
Effectiveness	High	High	High	High
Performance	Moderate	High	Moderate to High	Moderate
Scalability	High	Limited	Moderate to High	High

Cost Efficiency	Low or Moderate or High	Low to Moderate	Moderate	Moderate
Integration & Compatibility	High to Moderate	Low to Moderate	Moderate	High
Usability	Moderate	Moderate	Moderate	Low to Moderate
Best Use Cases	General-purpose or non-sensitive storage	Compliance-heavy sensitive storage	Mixed workloads	Global and resilient storage

(I) Public Cloud

Shared infrastructure provided by third-party vendors like AWS, Google Cloud, and Azure. Economies of scale and wide adoption.

Suitability for Data Storage Countermeasures:

- Strengths:
  - Encryption: Public clouds offer built-in encryption services, such as AWS Key Management Service (KMS) and Azure Encryption, ensuring data protection at rest and in transit.
  - Redundancy and Backup: Vendors provide robust backup and replication solutions across regions for disaster recovery.
  - Cost Efficiency: Pay-as-you-go models make public clouds cost-effective for data storage, especially for non-sensitive or bulk data.
  - Scalability: Ideal for organizations with growing or unpredictable storage needs.
- Challenges:
  - Limited control over physical infrastructure may pose challenges for compliance-heavy industries.
  - Shared tenancy can increase exposure to data leakage risks.
  - Somehow moderate evaluation for the Integration criterion, due to it enhances security and efficiency but is not the sole factor determining privacy. This leads also to a moderate performance.
- Challenges in Cost Efficiency (Moderate or High):
  - Cost efficiency is a significant challenge in implementing countermeasures for cloud data storage protection. Encryption-based solutions, such as homomorphic encryption and advanced key management systems, often incur **high costs** due to computational complexity, resource consumption, and compliance requirements. These methods demand substantial processing power, increasing cloud service expenses, particularly for large-scale data operations.
  - On the other hand, **moderate-cost** solutions like tokenization and basic encryption offer a balance between security and affordability. While they reduce storage and processing overhead, they still require infrastructure investment and ongoing maintenance. Additionally, compliance with data protection regulations (e.g.,

GDPR, HIPAA) adds indirect costs, including audits, updates, and specialized personnel.

- Organizations must weigh security needs against budget constraints, opting for a cost-effective mix of encryption, tokenization, and access controls. Finding the right balance ensures data protection without excessive financial strain, making cloud security both practical and sustainable for businesses.

Best Use Cases:

- Startups and SMEs requiring scalable and cost-effective storage.
- Storing less sensitive data and/or implementing robust encryption and access control mechanisms for sensitive information.

(II) Private Cloud

Dedicated infrastructure managed by the organization or a private vendor.

Suitability for Data Storage Countermeasures:

- Strengths:
  - Data Sovereignty: Ensures complete control over data, aiding compliance with regulations like GDPR and HIPAA.
  - Access Control: Organizations can implement granular access control tailored to their security needs.
  - Customization: High flexibility to design storage countermeasures, such as air-gapped backups or custom encryption protocols.
- Challenges:
  - High cost of infrastructure setup and maintenance.
  - Limited scalability compared to public clouds.
  - Low to moderate evaluation for cost efficiency, since security & privacy take priority over cost considerations.
  - Private cloud infrastructure requires high initial investment, reducing cost flexibility. Regulatory compliance & control needs often outweigh cost savings.

Best Use Cases:

- Industries with stringent regulatory requirements, such as healthcare, finance, and government.
- Organizations requiring full control over sensitive data.

(III) Hybrid Cloud

Combines public and private cloud infrastructure, enabling data and application movement between environments.

Suitability for Data Storage Countermeasures:

- Strengths:
    - Data Segregation: Sensitive data can be stored on private cloud infrastructure, while non-sensitive data resides on the public cloud, optimizing security and cost.
    - Scalability: Offers flexibility to handle peak loads by utilizing public cloud resources.
    - Backup and Disaster Recovery: Enhanced resiliency through cross-cloud backups.
- Challenges:
    - Complex integration of storage solutions across environments.
    - Potential latency issues in data synchronization, backup and recovery.

Best Use Cases:

- Enterprises need to balance security for sensitive data with scalability for non-sensitive workloads.
- Organizations requiring hybrid disaster recovery solutions.

(IV) Multi-Cloud

Utilization of multiple cloud providers to avoid vendors lock-in and ensure redundancy.

Suitability for Data Storage Countermeasures:

- Strengths:
    - Redundancy: Enhanced disaster recovery by distributing data across multiple providers.
    - Performance Optimization: Ability to select the best provider for specific workloads.
    - Risk Mitigation: Reduces dependency on a single vendor, improving security posture.
- Challenges:
    - Increased complexity in managing data across platforms.
    - Higher integration costs are due to differences in provider protocols and APIs.

Best Use Cases:

- Organizations with global operations require high availability and diverse compliance needs.
- Businesses prioritize resilience against provider outages or breaches.

The ideal cloud environment for data storage countermeasures depends on an organization's priorities:

- **Public Cloud**: Best for scalable and cost-effective storage solutions with encryption and redundancy for non-sensitive data.
- **Private Cloud**: Suitable for industries requiring strict compliance and full control over sensitive data.
- **Hybrid Cloud**: Offers a balanced approach for organizations with mixed workloads, combining security and scalability.

- **Multi-Cloud**: Ideal for businesses seeking resilience, redundancy, and vendor diversification, though it comes with higher complexity.

Each cloud environment provides distinct advantages for data storage countermeasures, allowing organizations to tailor solutions to their operational and regulatory needs.

5.2.2.5 Suitable Cloud Environments for Data privacy protection mechanisms countermeasures

Data privacy protection mechanisms aim to safeguard sensitive information from unauthorized access, breaches, and other privacy threats in cloud environments. These countermeasures include encryption, tokenization, access controls, anonymization, and monitoring solutions. The effectiveness of such mechanisms depends on the characteristics of different cloud environments—Public, Private, Hybrid, and Multi-Cloud—and their alignment with the organization's data privacy requirements. Below is an analysis of these environments for Data privacy protection mechanisms countermeasures (Table 17).

Table 17: Data privacy protection mechanisms category countermeasures for cloud environments based on key evaluation criteria

	Cloud environments			
Criteria	**Public Cloud**	**Private Cloud**	**Hybrid Cloud**	**Multi-Cloud**
Effectiveness	High	High	High	Moderate to High
Performance	Moderate	High	Moderate to High	Moderate
Scalability	High	Low to Moderate	High	High
Cost Efficiency	Moderate	Low to Moderate	Moderate	Moderate
Integration & Compatibility	High	Moderate	Moderate to High	Low to Moderate
Usability	Moderate	Moderate	Moderate	Moderate
Best Use Cases	General-purpose privacy	Stringent compliance, sensitive data	Mixed sensitivity workloads	High availability, global data

(I) Public Cloud

Shared infrastructure offered by providers like AWS, Google Cloud, and Azure. Focuses on scalability, availability, and cost efficiency.

Suitability for Data Privacy Protection Mechanisms:

- Strengths:
  - Access Controls: Providers implement multi-layered identity and access management (IAM) systems to restrict unauthorized access.
  - Anonymization: Services for anonymizing sensitive data, such as differential privacy techniques (Data masking, Data Tokenization), are available.
  - Monitoring: Public clouds offer native monitoring tools like AWS CloudTrail and Azure Security Center to track access and usage.
- Challenges:
  - Shared tenancy increases exposure to risks, such as side-channel attacks and data leakage, which may lead to affect the Performance and therefore the cost.
  - Dependence on the provider for compliance with data protection regulations like GDPR and HIPAA.

Best Use Cases:

- Organizations store moderately sensitive data with a focus on scalability and cost-effectiveness.
- Startups and SMEs that rely on provider-offered privacy tools.

(II) Private Cloud

Dedicated infrastructure is managed by the organization or a third party.

Suitability for Data Privacy Protection Mechanisms:

- Strengths:
  - Custom Encryption: Offers flexibility to deploy proprietary encryption or highly-secure standards and protocols.
  - Access and Identity Controls: Allows tailored implementations, and granular role-based access.
  - Data Sovereignty: Ensures full compliance with regional and industry-specific regulations by maintaining complete control over data.
  - Monitoring and Auditing: Organizations can deploy custom tools for privacy monitoring and data access audits.
- Challenges:
  - Higher costs for infrastructure setup and maintenance.
  - Requires in-house expertise to manage privacy countermeasures effectively.
  - Low to moderate evaluation for cost efficiency**,** since Security, privacy, and compliance take priority over cost considerations.
  - Private clouds require significant initial investment, making cost efficiency less flexible than in public clouds. Organizations prioritize control and customization, even if it leads to higher costs.

Best Use Cases:

- Enterprises in highly regulated industries, such as healthcare, banking, and government, handling highly sensitive data.
- Organizations require full control over privacy mechanisms and compliance processes.

(III) Hybrid Cloud

Combines public and private cloud infrastructures, enabling data and workload portability.

Suitability for Data Privacy Protection Mechanisms:

- Strengths:
  - Sensitive Data Segmentation: Enables storage of sensitive information on private clouds while using public clouds for general-purpose workloads.
  - Secure Workload Management: Balances privacy protection and operational scalability by assigning workloads based on sensitivity levels.
  - Integrated Monitoring: Cross-cloud monitoring and logging ensures consistent enforcement of privacy policies.
- Challenges:
  - Complexity in integrating privacy protection mechanisms across public and private environments.

- Potential latency in synchronizing privacy controls between environments. Moderate Usability and Cost efficiency.

Best Use Cases:

- Large organizations balancing security, cost efficiency, and scalability.
- Companies require secure backups for sensitive data.

(IV) Multi-Cloud

Utilizes multiple cloud providers to avoid vendor lock-in and enhance redundancy.

Suitability for Data Privacy Protection Mechanisms:

- Strengths:
  - Resilience: Improves privacy protections by reducing dependence on a single provider.
  - Customized Controls: Enables selection of the best provider for specific privacy needs (e.g., GDPR-compliant regions).
  - Monitoring Tools: Offers diversified tools for real-time access tracking and risk management across providers.
- Challenges:
  - Increased complexity in implementing consistent privacy controls across providers.
  - Some higher costs for integrating tools and ensuring compatibility.

Best Use Cases:

- Global enterprises manage data subject to diverse compliance requirements.
- Organizations prioritizing resilience and redundancy in data privacy.

The proper cloud environment for data privacy protection mechanisms depends on the nature of the data, the level of sensitivity, and organizational priorities:

1. **Public Cloud**: Best for scalable and cost-effective privacy solutions for non-critical data.
2. **Private Cloud**: Ideal for sensitive and compliance-critical data requiring tailored privacy mechanisms.
3. **Hybrid Cloud**: Suitable for organizations balancing secure storage for sensitive data and public cloud scalability.
4. **Multi-Cloud**: Effective for global enterprises requiring resilience, redundancy, and provider diversification.

Organizations must carefully assess their privacy requirements, data sensitivity, and compliance needs to choose the most appropriate cloud environment for their data privacy protection mechanisms.

5.2.3.6 Final comments and a proposed algorithm

Finally, we can propose the following 3 phases-algorithm for a company to *select* and finally *implement* the proper strategy for choosing the right countermeasures to protect cloud privacy from attacks, based on the above analysis. In the 1st phase we follow specific steps to choose the proper countermeasure category (or categories) and during the 2nd phase we follow specific steps to choose the specific countermeasures within the selected (from the 1st phase) category (or categories). Once the 2nd phase is over, the algorithm continues with the third phase to implement the countermeasures.

It is important here to note, that if a company has *budget restrictions* for its IT strategy and chooses the 'Cost Efficiency' as one of the desired criteria, the proposed algorithm will take these restrictions under consideration, during the whole process of the countermeasures' selection.

1st phase steps:

1) prioritize and select the criterion(a) (along with their desired values) which will satisfy the company's IT strategy, and it will base on them its budget for implementing the relevant countermeasures. Additionally, select (determine) the cloud environment on which it (or will) operates,

2) Check all tables 13-17 whether there is one or more countermeasure categories that satisfy the values given for the selected criteria with respect to the desired cloud environment. If no countermeasure converges, then the user should attempt to loosen the value for one or more of the selected criteria until it finds one or more converging countermeasure categories. In case the user reaches the situation that it cannot further loosen criteria values, it has to go back to step 1 and change the criteria given (along with their values),

2nd phase:

The goal of this phase is to select countermeasures from the picked categories in Step 1 based on their strengths, possible challenges and mainly on their values and justification, comparing them, as well. This maps to executing the following steps:

1) from tables 5, 7, 9-11 overview the countermeasures for every of the 1st phase converged category(ies). If there are more than one selected and converged categories, proceed to step 3 below,

2) based on the tables 5, 7, 9, 10 and 11 select the countermeasure(s), by weighting their pros and cons, values and best fit (even the budget for the IT strategy must take into account), and return it(them) to 1st phase. If you entail one or more common countermeasures within every converged category, then you must compare them, not only with their best use, strengths and treatable weaknesses within the current category they belong but also across all converged categories that they are involved. In each category, you compare all countermeasures that belong to it in order to select the best possible one. In case that one countermeasure prevails but it is also part of another converged category, then it needs to prevail also in that category in order to be selected.

In case we have a countermeasure that is selected in a converged category which, in turn, also belongs to another converged category which is not the best one, we have to choose: *a)* to select the specific countermeasure from the first category and the best one from the other converged category and combine them, or *b)* to select the unique best countermeasure from the converged

categories. The selection here, namely to select multiple countermeasures or only one, depends on additional factors and criteria, which must be taken into account, as the total cost, the usability, the easiness of the installation, the integration in the current infrastructure and generally, to use some or all of the remaining criteria not used at the beginning of the algorithm selection process. Regarding the proper classification and categorization of the countermeasures, we may rely on a risk assessment approach (see Section 5.2.3.7), taking into consideration that the countermeasures will also have to satisfy three main types of controls: preventive, detective and corrective[128].

3) return the selected countermeasure or the combination of them, to 3rd phase.

3rd Phase:

1) implement each from the selected countermeasure in Phase 2 (or the combination of them) on servers, on which you can test them, applying even '*mock penetration tests*', before launching them into production, and whether these are already available as a service by the cloud provider,

2) continue to control the performance of the realized countermeasure(s) to best tune and improve it(them).

As an example, for the above algorithm, let us consider a company (new start-up) which makes internet sales for individuals. It wants the protection of its clients' data (storing them in the cloud) and wants to spend a logical amount of money to assure cloud privacy. Also, it decides that it will use cloud computing services from two cloud providers to run its applications. Therefore, it considers its cloud environment as a *public-cloud environment*. The criteria on which it will base its IT strategy for cloud privacy are *usability* and *cost efficiency*, expressing as desired values, exactly 'Moderate", for the *usability* and for the *cost efficiency* the "Moderate" is the exact value accepted (filtered). *We have to note here, that the company has budget restrictions for its IT strategy and therefore it will realign the strategy and take the restrictions under consideration, during the whole process of the countermeasures' selection*.

We then continue our process/algorithm and check the content of Tables 13-17 in order to find out which countermeasure categories converge with respect to the public-cloud environment in terms of the desired criteria and their values. By inspecting these tables, we come into the following result:

Usability:	DLPs	x	IDS	x	MiTM	x	**Data Storage**	✓	**Data protection** ✓
Cost:	DLPs	✓	IDS	x	MiTM	x	**Data Storage**	x	**Data protection** ✓

In this respect, only the category *Data Protection* converges.

[128] *Preventive Controls* – Measures like encryption, access control, and multi-factor authentication prevent unauthorized access and data exposure.
Detective Controls – These include monitoring tools, intrusion detection systems, and audit logs to identify and alert on suspicious activities.
Corrective Controls – Incident response plans, backup recovery strategies, and automated security patching help mitigate the damage from security incidents.

Therefore, the “*Data protection*” category seems to be the most appropriate for achieving the security/privacy requirements of the organisation. Essentially, as this category applies to the public-cloud environment of the company, satisfying both criteria and with the desired tagging (evaluation values). Then, the company has to discuss, plan and suggest the best of the relevant countermeasures for this category (i.e. cryptography-oriented, probability-oriented, probabilistic hybrid, utility-preserving, anonymization, ranking-oriented, encryption and tokenization), benefiting of the strengths, considering some challenges, and gaining resilience and redundancy in data privacy (database efficiency and reduces data loss). Also, it should be selected, by taking into consideration their distinct features and performance as well as the kind of data that must be manipulated and the type of processing to be done on that data.

For the company, in consideration, and from the Table 11, a *Cryptography-oriented* type countermeasure or *Encryption-Tokenization* might be used, which will ensure privacy during processing and mitigating the risk of exposure, protect data privacy with confidentiality, securing storage, secrecy and transactions of clients’ data, offering strong privacy protection by preventing unauthorized access at any point in transit or storage (as the justification of our proposed countermeasures for this category).

A *cryptography-oriented countermeasure* secures data by using encryption algorithms such as AES or RSA. These methods ensure that even if data is intercepted, it remains unreadable without the decryption key. End-to-end encryption (E2EE) further enhances security by encrypting data before it leaves the user's device, preventing cloud providers or attackers from accessing sensitive information. This approach is crucial for compliance with data protection regulations like GDPR and HIPAA.

Alternatively, *encryption-tokenization countermeasures* replace sensitive data with non-sensitive placeholders (tokens) while storing the original information in a secure vault. Tokenization is particularly effective for protecting financial data, such as credit card numbers, as tokens are useless if stolen. Unlike encryption, tokenization does not require complex key management, reducing the risk of key exposure.

Both methods significantly improve cloud privacy, but the choice depends on the use case. Cryptographic encryption is ideal for securing large datasets and confidential communications, while tokenization is more effective for structured data protection in industries like banking and healthcare. For the company in question, a better choice for countermeasure is the **cryptography-oriented countermeasure.**

In general, and in the case of the involvement of a software developer, there is a need to implement a countermeasure, or in the case of exploiting existing countermeasure services, a different role might be needed by the organization, before the development of the integration of them.

It must be noted that much attention should be paid, when data moves between different cloud environments (private-to-public, multi-cloud, or hybrid cloud setups), then the countermeasures must be integrated and work together to maintain security, privacy, and compliance. Integrating countermeasures across clouds requires a well-structured approach, indicatively, focusing on encryption, IAM, DLP, compliance, secure transfer methods., automation,

Zero Trust principles, and AI-driven security tools, which in combination can enhance cooperation and reduce security gaps when data moves between clouds.

In conclusion, a company should examine, on which *criterion(a)* will rely on (as the most important for them) and on which *cloud environments* will (it) operate(s), in order to select the *best category countermeasures* and therefore to develop its IT strategy with the proper countermeasures, taking under consideration for every specific countermeasure its effectiveness, challenges and implementation procedure, so as finally to protect the cloud privacy for threats and attacks.

Initially, the algorithm does a look on the countermeasure landscape at a higher-level of abstraction, attempting to select the most appropriate category of countermeasures. It proposes the category or categories on which most of the focus should be.

Then, companies should proceed to the 2nd phase of the algorithm in a multi-level security strategy/architecture that integrates the countermeasures of the selected category or combines countermeasures from different selected categories when they are common.

5.2.3.7 Holistic Risk-Approach Strategy for Cloud Privacy Protection

In today’s interconnected world, cloud computing has become integral for businesses. However, the widespread adoption of cloud services has led to an increased focus on protecting sensitive data stored in the cloud. Protecting cloud privacy requires a systematic risk approach strategy that not only addresses existing vulnerabilities but also prepares for potential future threats. For this reason, we finally propose a *risk-approach strategy* which is based on specific critical stages: (a) collecting previous or possible threats, attacks, and vulnerabilities, (b) rank countermeasures to mitigate these risks based on their effectiveness to address the respective critical risks discovered, applying also the previously explained (section 5.2.2.6) algorithm, for the selection of the countermeasures. This approach can be also considered as an enrichment in the algorithm. More specifically:

Stage 1: Collecting risks from possible or previous Threats, Attacks, and Vulnerabilities

The first stage in any comprehensive risk approach to cloud privacy protection is understanding the historical context of cyber threats, attacks, and vulnerabilities. This phase helps businesses identify recurring patterns and understand the specific risks associated with their cloud infrastructure. Important steps to perform:

1. **Threat Intelligence Gathering:** A robust threat intelligence system should be implemented to continuously gather data on emerging risks. This includes subscribing to threat intelligence feeds, utilizing industry reports, and engaging with cybersecurity communities. By monitoring known threat actors and attack patterns, a company can gain insight into how their cloud infrastructure may be targeted.
2. **Historical Incident Analysis:** The company should assess past incidents of security breaches, attacks, and data leaks within its own organization and the industry at large. This can involve reviewing historical incident reports, conducting internal post-mortem analyses, and learning from the mistakes of others. For instance, past incidents of data breaches in the

cloud, like misconfigured cloud storage buckets or weak authentication methods, can highlight specific risks the company must address.

3. **Vulnerability and Risk Assessment:** Cloud infrastructures often have specific vulnerabilities, such as weak access controls, insufficient encryption, or insecure APIs. A thorough vulnerability assessment should be conducted regularly to identify potential weaknesses in cloud architecture, network configurations, and third-party services. By keeping track of past vulnerabilities and patching known issues, the company can build a stronger security posture for future operations. Once threats are identified, organizations assess the impact and likelihood of these threats affecting cloud data. This involves evaluating existing security controls and identifying gaps.
4. **Regulatory Compliance Review:** Compliance regulations, such as GDPR, HIPAA, and CCPA, influence how cloud data should be handled. A review of previous compliance violations, audit findings, or legal issues can help the company understand how well it has been adhering to regulatory standards and where further improvements are needed.

Stage 2: Risks prioritization, ranking of Countermeasures and the use of the selection algorithm

Once the risks are prioritized and measured (using their probability of occurrence and their impacts) by company and vulnerabilities are identified, this stage involves the ranking of the appropriate countermeasures to mitigate these threats (or any potential). The selected countermeasures originate from the proposed selection algorithm (from 2nd phase).

Then, we can either find one countermeasure that addresses all previously identified risks or a minimum combination of countermeasures that addresses these risks. Once we find one or more combinations of countermeasures that address the risks, we can then proceed to select one of them as the best. In that case, we use the algorithm, which considers further criteria, apart from effectiveness.

Finally, we have the risks to confront, the criteria to optimize and the countermeasures to select. Then we need to find the best possible combination of countermeasures that completely cover all risks and satisfy the criteria given.

Conclusion

A robust risk approach strategy for cloud privacy protection requires a methodical process of understanding past threats, attacks, and vulnerabilities and then implementing tailored countermeasures. By being informed about emerging risks, categorizing countermeasures based on effectiveness, and continuously monitoring the cloud environment, companies can build a strong defense against potential threats and ensure the integrity and privacy of their cloud-based data. Regular assessment and adaptation of these strategies will ensure the company remains resilient against the ever-evolving landscape of cybersecurity threats.

6 *Further research*

In this chapter we present some future research directions on enhancing cloud computing privacy. We focus on presenting challenges for cloud security and privacy as well as trends and promising research directions to cover them. Moreover, we present recent trends and future directions for cloud computing that can positively impact the development of new or the extension of existing countermeasures.

6.1 Future research on cloud privacy and countermeasures

6.1.1 New approaches and research for DLPs

Data leakage is an ongoing problem in the field of information security. Both academics and practitioners are continuously working towards developing data leakage prevention and detection methods to mitigate this problem. DLPs are increasingly recognized as the preferred solutions for identifying, monitoring and protecting confidential data.

From industry references and the literature, it is evident that data leakage detection and prevention methods are inadequately studied (Alneyadi, S., et al, 2014). Moreover, most of the current methods suffer from serious limitations, especially when confidential data evolve. This is because they mainly depend on inflexible techniques. Even with some robust fuzzy fingerprinting and statistical analysis, confidential data semantics can be leaked using various obfuscations (techniques used to obfuscate data usually are encryption, tokenization, and data masking). Therefore, a potential research question in this area is “how to detect semantically the content of confidential data in order to prevent data leakage”.

Thus, effective future DLPs should have the ability to classify confidential data semantically even if such data evolve. In particular, DLPs should have the ability to heuristically detect such data without the need for managing exact copies of existing and new data. In addition, detection techniques require extensive analysis, which includes deep content inspection and comprehensive indexing. This can impose a serious challenge to DLPs as they could exhaust the available computational and storage capabilities. Hence, efficient DLPs should be able to perform the “heavy” detection tasks with minimal computational and storage requirements.

6.1.2 Future research directions for IDS

Statistical approaches are commonly employed for anomaly detection in IDS[129] (Schmidt, F., 2018). These techniques involve the use of statistical methods to establish normal behavior baselines and detect deviations from these baselines. Outlier detection algorithms, such as the statistical outlier detection method or the Z-score method, are used to identify data points that significantly deviate from the expected behavior (Zhang, X. et al, 2019)[130]. Time series analysis techniques, such as autoregressive integrated moving average (ARIMA) models, are used to detect anomalies in temporal data. Statistical modeling approaches, such as Gaussian mixture models or hidden Markov models, are utilized to capture the statistical characteristics of normal behavior and detect anomalies based on deviations from the learned models.

Hybrid IDS approaches combine both statistical and machine learning[131] (ML) techniques to improve the accuracy and effectiveness of anomaly detection in IDS (Sajid, M., et al, 2024)[132]. By leveraging the strengths of different approaches, hybrid models can provide enhanced detection capabilities. For example, a hybrid approach may use statistical techniques to establish baseline behavior and machine learning algorithms to classify instances as normal or anomalous. If the baseline behavior is established, then it is computationally rapid to classify instances as normal or anomalous. Then, we will try to make more complex solutions in order to get a better result i.e, for selecting and applying in conjunction ML algorithms. They would require their own training and thus rely on different models. This combination allows for a more comprehensive and robust anomaly detection system.

The domain-specific and characteristic challenges of applying Machine Learning (ML) techniques for developing an IDS is an important topic. For effectively applying ML techniques, issues reside in the proposition of anomaly detection that does not coincide with the fundamental theory of ML techniques. ML techniques are known for their learning ability and adapting the knowledge to make predictions for the given data. There are studies (Thakkar A. and Lohiya R., 2022[133]) which discuss the generic process of ML techniques applied for intrusion detection and classification. The process consists of four phases, namely, data collection, data pre-processing, feature engineering, and classification. ML methods, when applied to the anomaly detection dataset, extract and learn the patterns from the features and predict the attack.

The ML techniques are used for building IDS, with challenges such as: (i) high false alarm rate; (ii) lack of training data; (iii) real-time intrusion detection; (iv) need for alert correlation; (v) proficiency in handling encrypted network traffic; (vi) Assessing potential risks; (vii) challenge in

---

[129] F. Schmidt, F. Suri-Payer, A. Gulenko, M. Wallschläger, A. Acker and O. Kao, "*Unsupervised Anomaly Event Detection for Cloud Monitoring Using Online Arima*," 2018 IEEE/ACM International Conference on Utility and Cloud Computing Companion (UCC Companion), Zurich, Switzerland, 2018, pp. 71-76, doi: 10.1109/UCC-Companion.2018.00037.

[130] X. Zhang, L. Liu, X. Chen, S. Xie and L. Lei, "*A Novel Multitemporal Cloud and Cloud Shadow Detection Method Using the Integrated Cloud Z-Scores Model,*" in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 12, no. 1, pp. 123-134, Jan. 2019, doi: 10.1109/JSTARS.2018.2889150

[131] As AI and ML become more integrated into cybersecurity, attackers are finding ways to manipulate these technologies. Adversarial attacks on AI involve feeding misleading data into machine learning models to cause incorrect decisions.

[132] Sajid, M., Malik, K.R., Almogren, A. et al. "*Enhancing intrusion detection: a hybrid machine and deep learning approach".* J Cloud Comp 13, 123 (2024). https://doi.org/10.1186/s13677-024-00685-x

[133] Ankit Thakkar, Ritika Lohiya, "*A Review on Challenges and Future Research Directions for Machine Learning-Based Intrusion Detection System"*, Vol.:(0123456789)1 3Archives of Computational Methods in Engineering (2023) 30:4245–4269 https://doi.org/10.1007/s11831-023-09943-8, p.4265

handling intrusion evasion attacks; (viii) securing industrial control systems. Based on these challenges, three are potential future research directions that can be explored to strengthen the research performed in the area of ML-based IDS. Moreover, the intensity of the network attack accomplishment is increased day by day. Therefore, there is a need to build a model with advanced techniques that can predict and classify the attacks, assess the risk associated with the attacks, aggregate the alerts, reduce the number of false alarm rates, and have a brisk response for intrusions detected within a given time frame.

6.1.3 Cloud security trends in the future

In the fluid world of cloud security, regulatory compliance is an evolving objective. Staying ahead means more than just complying with existing regulations; it involves anticipating and gearing up for upcoming developments & evolutions. This forward-thinking approach to compliance positions organizations as rule followers and as leaders in security excellence.

Adapting to regulatory changes requires a deep understanding of the letter and the spirit of the law. It is about integrating compliance into the very fabric of an organization’s cloud security strategy. This integration allows companies to pivot quickly and efficiently when new regulations come into play, minimizing disruptions and maintaining a continuous state of compliance.

Staying ahead of regulatory changes is more than a legal necessity; it is a competitive advantage. It demonstrates to customers and stakeholders that an organization is committed to maintaining the highest data protection and privacy standards. By adopting this approach, regulatory compliance solidifies its role as a fundamental element in establishing trust and legitimacy in the online realm.

The cloud security landscape is constantly evolving, with new technologies and methodologies emerging to address changing cloud data security challenges. As we head further into 2024, several key cloud computing trends have started shaping the future of cloud security. Some of them are presented below:

1. Widespread adoption of quantum-resistant encryption. With the looming threat of quantum computing's ability to break traditional encryption methods, the demand for quantum-resistant encryption algorithms has skyrocketed. In 2024, we are witnessing a widespread adoption of post-quantum cryptography (PQC) algorithms across cloud service providers and enterprises to ensure long-term data security and confidentiality in cloud environments. With the potential for quantum computing to break current encryption standards, PQC is gaining traction. These cryptographic methods aim to resist quantum-based attacks and ensure the futureproofing of cloud storage solutions.

2. Zero-Trust Architecture. The traditional perimeter-based security model is becoming increasingly obsolete in the cloud era. Zero-trust architecture, which assumes no user or device should be trusted by default, has emerged as the new gold standard for cloud security. This approach involves continuous verification of user identities, devices, and access privileges, providing a more robust and granular level of security. For example, some organizations are implementing innovative MFA (Multi Factor Authentication) approaches for accessing their cloud infrastructure, such as requiring QR codes for registration or even giving out special verification hardware/software for accessing physical servers (ie: Time-Based One-Time Password (TOTP), SMS text message token,

Email token, Hardware security key, Biometric authentication, Security questions, Risk-based authentication).

3. Shift towards unified security management. With the proliferation of cloud services and the increasing complexity of cloud architectures, managing security across multiple cloud environments (even hybrid ones) has become daunting. In response, unified cloud security management platforms are emerging as Microsoft's unified security operations platform provides a single platform for end-to-end security operations (*SecOps*). It integrates security information and event management (SIEM), security orchestration, automation, and response (SOAR), extended detection and response (XDR), posture and exposure management, cloud security, threat intelligence, and generative AI solutions.

4. Emphasis on cloud security automation and orchestration. As cloud environments become more dynamic and complex, manual security processes are proving increasingly inefficient and prone to human error. In 2024, we are witnessing a growing emphasis on security automation and orchestration, leveraging artificial intelligence and machine learning to automate routine security tasks, respond to threats in real time, and streamline incident response processes.

6.1.4 Cloud privacy protection advancements

Recent advancements in cloud privacy for data storage focus on innovative encryption techniques, privacy-enhancing technologies, and the preparation for quantum computing. Here are some key developments:

Fully Homomorphic Encryption (FHE):

FHE enables computations on encrypted data without decryption, ensuring data privacy throughout its lifecycle—during storage, transmission, and processing. While FHE offers robust security, challenges like longer execution times are being addressed with advancements in hardware and software optimization. This technology is seen as a transformative tool for securing sensitive data in untrusted cloud environments.

Privacy-Enhancing Technologies (PETs):

Methods like differential privacy, federated learning, and secure multi-party computation are increasingly adopted to balance privacy with the need for data analysis. These techniques allow organizations to derive insights from aggregated data without exposing individual records, which is particularly important in sectors like healthcare and finance.

Blockchain Integration for Privacy:

Blockchain is being used to enhance data privacy through secure, decentralized storage and access control mechanisms. By leveraging immutable and transparent record-keeping, blockchain ensures better accountability while preventing unauthorized access

AI in Privacy Management:

Artificial intelligence is utilized to identify vulnerabilities and automate privacy compliance processes. AI-driven tools help classify sensitive data, detect anomalies, and enforce privacy protocols in real-time, reducing human error in data protection

These advancements aim to address the dual challenges of securing sensitive information and maintaining compliance with evolving privacy regulations like GDPR and CCPA, ensuring the cloud remains a trustworthy platform for data storage and processing.

6.2 Trends and future directions for cloud computing

Leaders across many industries need to pay attention to emerging trends in both cloud computing and cloud privacy, as these innovations are transforming everything from business operations to data management, digital services and more. Many research trends are also based on *Artificial intelligence (AI)* and *machine learning*, which lead the charge in the fight against cloud computing threats, positioning themselves as pivotal elements in this enduring battle. Far from being mere tools, they represent the architectural elements of a new era in cybersecurity. These technologies continuously learn and evolve, sharpening their skills to outwit the most sophisticated digital adversaries. Imagine a security system that learns from every attack and anticipates future threats based on patterns and anomalies. These technologies/future trends are analysed as follows.

i) Serverless Computing

Serverless computing is an emerging trend in cloud computing that allows developers to build and run applications without managing the respective server infrastructure. Key benefits include cost efficiency through a pay-per-use model, simplified operations as cloud providers handle infrastructure, automatic scaling to meet demand and suitability for event-driven architecture. By reducing complexity and supporting a microservices architecture, it enhances development agility and accelerates time-to-market.

Serverless security, at first glance one could argue that serverless computing is intrinsically more secure than its predecessors because of its characteristics (e.g., the short duration of functions), or due to the fact that it could inherit security features already developed for other virtualization solutions (Marin E, Perino D., 2022). Yet, serverless computing brings many new, idiosyncratic security challenges that open the door for new types of security attacks. Further, implementing serverless applications requires a major change in mindset from software developers, not only in the way applications are written but also in the way they are protected from security attacks. These latter requirements are rarely met, hence introducing new vulnerabilities. However, with the increase in volume and diversity of attacks against the cloud, it becomes apparent that security and privacy will be a key factor which, if not addressed, could hamper the widespread adoption of serverless computing.

ii) Quantum Cloud Computing

Integrating quantum computing into cloud platforms promises exponential growth in processing power, revolutionizing complex calculations and encryption methods. This synergy will catalyze cryptography, drug discovery and financial modeling advancements by offering unparalleled computational capabilities. Quantum cloud computing represents a disruptive shift, enabling faster problem-solving and unlocking innovation and scientific research possibilities across industries. Quantum cloud computing represents a groundbreaking fusion of quantum mechanics and cloud technology. It leverages the principles of quantum physics to foster the development of new computational methodologies that surpass the capabilities of classical computing. One of the most promising aspects of quantum cloud computing is its potential to enhance data security and privacy

significantly. Quantum systems enable what is known as quantum homomorphic encryption, which allows encrypted data to be processed without needing to decrypt it first. This ensures that sensitive information remains secure and private, even while being processed in the cloud.

ii) AI-Driven Cloud Management

AI-driven cloud management is revolutionizing how cloud resources are optimized and maintained. Through predictive analytics and automation, AI can dynamically allocate resources, anticipate failures and manage workloads more efficiently. This not only reduces operational costs but also enhances performance and reliability. As cloud environments become more complex, AI's role in managing these systems is becoming indispensable for achieving scalable and resilient cloud infrastructures. By analyzing user behavior, AI can detect anomalies that may indicate malicious activity from within an organization. Unusual access patterns, large data transfers and other suspicious behaviors can trigger alerts, enabling security teams to investigate and respond before significant damage occurs.

There are software tools (like Orca Security) which uses advanced analytics and machine learning algorithms to rapidly analyze and detect anomalies in cloud feeds, workloads, and configurations that might indicate malicious activity.

iii) Confidential Computing

Confidential computing is emerging as a transformative trend in cloud computing, providing advanced data protection through hardware-based Trusted Execution Environments (TEEs). Isolating sensitive data and workloads ensures secure processing even in untrusted environments. This technology is crucial for industries with stringent data privacy requirements, such as finance and healthcare, enabling them to leverage cloud capabilities without compromising security. It provides a novel, clearly defined security boundary, isolating sensitive data within trusted execution environments during computation.

This means services can be designed to segment data based on the least-privilege access principles, while all other code in the system sees only encrypted data. Crucially, isolation is rooted in novel hardware primitives, effectively rendering even cloud-hosting infrastructure and its administrator’s inability to access the data. This approach creates more resilient systems capable of withstanding increasingly sophisticated cyber threats, thereby reinforcing data protection and sovereignty in an unprecedented manner.

iv) Hybrid Multi-Cloud

Hybrid multi-cloud combining private cloud with public cloud is emerging. Companies can have sensitive data securely in private cloud and use public cloud to solve scalability issues. It will help to optimize cost by using public clouds for ad-hoc needs. Disaster recovery will be improved by redundancy across cloud environments. A recent report compiled by Digital Guardian showed that 50% of companies using public or hybrid clouds report losing data. Three-quarters of those respondents say up to 20& of their data stored in the cloud is not appropriately protected.

Recent studies (Zardari M.A. et. al,. 2013[134]) suggested a HMCDS (Hybrid Multi Cloud Data Security model and Data classification) model which is hybrid of public and private clouds.

[134] M. A. Zardari, L. T. Jung and M. N. B. Zakaria, "*Hybrid Multi-cloud Data Security (HMCDS) Model and Data Classification,*" 2013 International Conference on Advanced Computer Science Applications and Technologies, Kuching, Malaysia, 2013, pp. 166-171, doi: 10.1109/ACSAT.2013.40.

HMCDS improves the efficiency of data retrieval, data confidentiality and availability in cloud computing. This model provides two levels of data confidentiality, i.e., high level and low level, and also provides efficient data retrieval. The high level of confidentiality is achieved by implementing HMCDS model with multi-clusters, and the low level of confidentiality is achieved using data classification technique in cloud business model. The HMCDS model provides data confidentiality to the customers by using a hybrid multi-cloud with multi-clustering technique. The classified data is stored into the public cloud (here, we considered three cluster clouds, and each cluster contains three inner clouds).

v) Generative AI

The future of the cloud is all about the industry cloud where more and more functions are being developed for industry-specific use cases like banking, insurance, healthcare, media, and so on. One disrupting trend or technology shaping the future is Generative AI. The next wave of cloud is generative AI-infused industry cloud. This will have two merging facets: The first is applying generative AI to industry use cases through cloud LLM/ SLMs, and the second is applying cloud LLMs to engineer and manage the generative AI industry workloads on the cloud.

Generative AI transforms how we approach security in cloud environments by enabling proactive, adaptive, and intelligent security measures. We are entering a new era of cloud security by leveraging AI's ability to process vast amounts of data, detect vulnerabilities in real-time, and adapt to emerging threats. Generative AI's predictive capabilities are among its most significant advantages. By learning from historical data, AI models can forecast potential security vulnerabilities before they become exploitable weaknesses. While generative AI offers immense benefits, it also introduces new challenges regarding data privacy. AI systems often require large amounts of data for training, raising concerns about privacy and security. Organizations must ensure that sensitive data is encrypted and handled according to strict privacy regulations (GDPR). To mitigate these risks, data anonymization, encryption, and secure multi-party computation can be employed during AI training processes. Responsible AI deployment means safeguarding data while still leveraging AI's power for enhanced cloud security.

Consider that the cybersecurity landscape is like an enormous network where information sharing becomes the cornerstone of the defense strategy. By clubbing knowledge, resources, and expertise, industry may elude cybercriminals. Further, innovation is promoted by the resultant mixture of different perspectives and expertise combined to achieve a solution to complex problems related to security. It is a synergy wherein each participant, from tech giants to startups, becomes an integral contributor towards building a strong security posture.

The collaboration goes right into the public domain, with public entities and regulatory bodies. Such a partnership is, in fact, very much necessary when developing and implementing standards and practices serving the collective interest. By cooperation, the industry can thread through all the intricacies of cloud security more effectively, offering a much safer digital future for all.

7 Conclusions

In this final chapter, we first present the real contribution of this thesis in the research of the cloud privacy vulnerabilities, threats and attacks plus relevant countermeasures as well as its main benefits. Then, we explicate the personal experience that has been attained by conducting this thesis. We, finally, present the major challenges in cloud security and privacy as well as the critical research directions to follow.

7.1 Thesis contribution and benefits

This research work is motivated above all by the need to conduct extensive studies and determine the possible attacks in the cloud computing environment, together with assessing the probable impacts on cloud services. With the gradual but complete paradigm shift that has taken place towards cloud computing, it has now been amazingly realized that new avenues must be pursued to aptly support and provide the required functionalities sought after by today's business world. However, everything is said and done, this migration and transition towards cloud-based computing is not all smooth sailing and safe waters-as a matter of fact, it opens a floodgate of new vulnerabilities and serious security threats for the unwary adopter of such technologies.

Also, this research aims to contribute to cloud privacy by analyzing and categorizing the relevant attacks, threats, vulnerabilities and countermeasures, developing a set of criteria to classify the countermeasures categories, and proposing a 3-phase algorithm that enables us to choose the right countermeasures from specific categories based on the organisation’s requirements on the suggested criteria and the type of cloud that it intends to use or already exploits. Finally, we have explained the main challenges in cloud security and privacy as well as presented research trends and future work directions to address them.

More specifically, we developed, initially, a novel model for the classification and taxonomy of cloud privacy attacks, encompassing both traditional and emerging threats. We also conducted a rigorous analysis of the root causes and potential impacts of these attacks on sensitive data stored and processed in cloud environments.

Then, we compared the countermeasure categories in which we have classified existing countermeasures, based on selected key evaluation criteria by taking also into account the different cloud environments in which an entity (e.g., an organization) might operate. As a result, each countermeasure category was assigned with a particular assessment value (High, Low or Moderate) for each criterion and possible cloud environment.

We, then, constructed an innovative algorithm with discrete steps to follow, to finally select the best-fit countermeasures for protecting cloud privacy and security. A private company can use

this tailor-made methodology to empower and protect its cloud data. Similarly, a public organization can use this algorithm and methodology to include it in the technical specifications for every procurement contest and invitation launches, related to its cloud developments (for security and privacy).

7.2 Experience gained

The research aimed at attaining a deep understanding of cloud computing concepts as it assisted in developing a strong foundation in cloud computing models, architectures, and services. The skills in conducting security assessments and vulnerability analysis were also well developed plus the ability to identify and assess potential vulnerabilities in cloud environments.

As an additional experience, we gained the ability to conduct a literature review and stay updated on the latest research. This was achieved by developing the skills to identify relevant research papers, critically analyze them, and synthesize the respective findings. At the same time, the capacity to identify research gaps and formulate research questions was increased. In particular, the thesis aimed, also, at learning how to identify areas where further research is needed and how to frame research questions to address those gaps.

Very important was the ability to think critically about complex security and privacy challenges, since it developed the skills to analyze complex problems, and evaluate the feasibility of their solutions. The capacity to adapt to new technologies and security threats played an important role, by learning to stay updated on the latest trends in cloud computing and cybersecurity and to adapt to our research to address emerging threats.

7.3 Major challenges and critical research directions to follow

Highlighting different research challenges with respect to the cloud service models which supply capabilities and satisfy requirements, will focus initially to application security.

Typically, cloud-based applications are delivered through the Internet to the users for fulfilling their requirements. Therefore, any vulnerability in the traditional web application model is automatically inherited by cloud applications. Cloud users can access their data through enterprise-distributed (i.e., multi-tenant) cloud applications.

The security and availability of these applications are highly dependent on the behavior and quality of these applications and of the cloud services that they encompass. One way is to encrypt the outsourced data for confidentiality and security purposes. Moreover, the security of these applications also depends on APIs. These software interfaces or APIs are the external access points for cloud services that need to be properly protected and secured. The existence of several virtualization technologies can also affect the security of these applications.

In an era where cloud environments are becoming increasingly complex, we need technologies and practices to act as the guiding principles ensuring unwavering security. The essence of *container security* lies in its focus on safeguarding each part of the cloud environment, like

securing every room in a building. Container security technology refers to the tools, practices, and methodologies used to protect containerized applications, their infrastructure, and the data they handle from potential threats. Containers are lightweight, portable units used to package applications and their dependencies, making them popular for modern software development. Container security technology is closely related to cloud environments because containers are frequently deployed in cloud-based infrastructures, whether public, private, or hybrid. They are interconnected, because:

Cloud is the Primary Host for Containers, the cloud security is a shared responsibility, Containers in the cloud need to be protected against these, container security technologies often integrate with cloud-native security tools, cloud environments often host Continuous Integration and Continuous Deployment (CI/CD) pipelines to build, test, and deploy containerized applications, Orchestrators are often deployed in cloud environments, and container security technology in the cloud ensuring workloads meet compliance requirements (GDPR).

Understanding the security and privacy risks in cloud computing and developing efficient and effective solutions are critical for its success (Takabi, D. et. al., 2011). Although clouds allow customers to avoid start-up costs, reduce operating costs, and increase their agility by immediately acquiring services and infrastructural resources when needed, their unique architectural features also raise various security and privacy concerns.

Cloud computing environments are multidomain environments in which each domain can use different security, privacy, and trust capabilities and potentially employ various mechanisms, interfaces, and semantics. Such domains could represent individually enabled services or other infrastructural or application components. Service-oriented architecture is naturally relevant technology to facilitate such multi-domain formation through service composition and orchestration. It is important to leverage existing research on multidomain policy integration and the secure-service composition to build a comprehensive policy-based management framework in cloud computing environment.

Looking ahead, one certainty stands out: the terrain of cloud security is continuously evolving. It demands direct actions like vigilance, innovation, and a deep understanding of the challenges and solutions that shape it. By staying informed and adaptable, organizations can navigate this landscape as participants and pioneers.

At its core, cloud security does not map just to technicalities and complex systems; it encompasses a human story, often unnoticed amidst technological aspects. This narrative is woven from the dedication and perseverance of professionals who are the sentinels of digital integrity. These individuals are silent guardians, diligently working out of sight to maintain the security of digital assets.

This human aspect brings emotion and reliability to the otherwise technical realm of cloud security. It is about the commitment of teams working round the clock to thwart cyber threats, the resolve of organizations to protect their customer’s data, and the collective growth of an industry that never stops learning and evolving. Understanding this human dimension adds depth to the conversation about cloud security. It is a reminder that at the end of every strategy, every line of code, and every innovation are individuals whose efforts and ingenuity continue to shape a safer digital world. It depends, ultimately, on humans to realize the future needs and future work

directions for cloud security and privacy and to act with a manner using and discovering technological tools to empower cloud technology for the good of society.

Finally, the authors’ opinion lies on the necessity of the close following the research trends which are based on *Artificial intelligence (AI)*, *machine learning* and other techniques (Generative AI, quantum computing, etc), which lead to the fight against cloud computing threats to build new architectural schemes for cybersecurity, in combination with a strategy to continuous find the best approach for cloud attacks threating privacy. Cybersecurity breaches are not limited to cloud security failures or privacy violations. Attacks on supply chains, IoT devices, critical infrastructure, and financial systems pose significant risks that require robust defense mechanisms. As technology evolves, so do cyber threats, making it essential for organizations to stay ahead with proactive security measures. Understanding these diverse cybersecurity breaches helps businesses and individuals better protect their assets and operations from emerging threats.

We must realize that as technology improves in the area of countermeasures for fighting cloud privacy attacks, at the same time, new kinds and methods are invented by hackers. Therefore, as technology changes, we must follow it, but we have also to keep in mind that we must always rely on a structured strategy using the optimum combination of cloud privacy countermeasures, based on various parameters, as we have presented in our thesis algorithm.

References

1. Ahmed M, Litchfield AT (2018), "*Taxonomy for identification of security issues in cloud computing environments*", J Comput Inf Syst 58(1):79–88
2. Almomani, A., et al., 2015. "*A survey of phishing email filtering techniques*". Commun. Surv. Tutor. IEEE 15 (4), 2070–2090.
3. Almorsy, M., Grundy, J.C., & Müller, I. (2016). "*An Analysis of the Cloud Computing Security Problem*". ArXiv, abs/1609.01107.
4. Alneyadi S., Sithirasenan E., Muthukkumarasamy V., "*A survey on data leakage prevention systems*", Journal of Network and Computer Applications, 2016, Elsevier, http://dx.doi.org/10.1016/j.jnca.2016.01.008
5. Amara, N., Zhiqui, H., Ali, A., "*Cloud Computing Security Threats and Attacks with their Mitigation Techniques*", 2017 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery, 978-1-5386-2209-4/17© 2017 IEEEDOI 10.1109/CyberC.2017.37244
6. Araújo, J.D., Abdelouahab, Z., 2015. "*Virtualization in intrusion detection systems: a study on different approaches for cloud computing environments*". IJCSNS 12 (11), 10.
7. Bakshi Aman, Dujodwala Yogesh B., "*Securing cloud from DDoS Attacks using Intrusion Detection System in Virtual Machine*" ICCSN '10 Proceeding of the 2010 Second International Conference on Communication Software and networks, pp. 260-264, 2010, IEEE Computer Society, USA, 2010. ISBN: 978-0-7695-3961-4.
8. Basu S et al (2018) "*Cloud computing security challenges and solutions—a survey*". In: Proceedings of the IEEE 8th Annual on Computing and Communication Workshop and Conference (CCWC), pp 347–356
9. Bhadauria R., Sanyal, S. (2014), "*Survey on Security Issues in Cloud Computing and Associated Mitigation Techniques*", International Journal of Computer Applications, April 2012, DOI: 10.5120/7292-0578 · Source: arXiv, New York, USA.
10. Bhadauria, R., et al., 2014. "*A survey on security issues in cloud computing*", arXiv preprint arXiv:1109.5388, 2014
11. Bouayad, A., et al., 2015. "*Cloud computing: security challenges*",. In: Proceedings of the Information Science and Technology (CIST), 2012 Colloquium in. 2015. IEEE.
12. Brocke, J.V., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., Cleven, A.: "*Reconstructing the giant: on the importance of rigor in documenting the literature search process*". Presented at the http://www.alexandria.unisg.ch/Publikationen/67910 June 10 (2009)
13. Catteddu, D. (2014). "*Cloud Computing: Benefits, Risks and Recommendations for Information Security*". In: Serrão, C., Aguilera Díaz, V., Cerullo, F. (eds) Web Application Security. IBWAS 2009. Communications in Computer and Information Science, vol 72.
14. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-16120-9_9). ENISA report
15. Council, S.-a.-a.-S.E., 2016. "*Software-as-a-Service; A Comprehensive Look at the Total Cost of Ownership of Software Applications*."
16. Dzombeta S, Stantchev V, Colomo-Palacios R, Brandis K, Haufe K (2014), "Governance *of cloud computing services for the life sciences*", IT Prof 16(4):30–37
17. Duncan, A.J., Creese, S., Goldsmith, M., 2012. "*Insider attacks in cloud computing*". In: Proceedings of the Trust, Security and Privacy in Computing and Communications (TrustCom), 2012 IEEE 11th International Conference on. 2015. IEEE

18. Fernandes, D.A., Soares, L.F., Gomes, J.V., Freire, M.M. and Inácio, P.R. (2014), "*Security Issues in Cloud Environments: A Survey*". International Journal of Information Security, 13, 113-170. http://dx.doi.org/10.1007/s10207-013-0208-7
19. Gentry C (2009) May, "*Fully homomorphic encryption using ideal lattices*", STOC 9:169–178
20. Ghorbel, A., Ghorbel, M., Jmaiel, M., "*Privacy in cloud computing environments: a survey and research challenges*", Published online: 23 January 2017, Springer Science+Business Media New York 2017 Springer Science+Business Media New York 2017, DOI: 10.1007/s11227-016-1953-y
21. Grance, T. and Jansen, W. (2018), "*Guidelines on Security and Privacy in Public Cloud Computing*", Special Publication (NIST SP), National Institute of Standards and Technology, Gaithersburg, MD, [online], https://doi.org/10.6028/NIST.SP.800-144
22. Hashizume, K., Rosado, D.G., Fernández-Medina, E. et al. "*An analysis of security issues for cloud computing*". J Internet Serv Appl 4, 5 (2015). https://doi.org/10.1186/1869-0238-4-5
23. Hussain S., Fatima M., Saeed A., Raza I., Shahzad R., "*Multilevel classification of security concerns in cloud computing*", Applied Computing and Informatics, Volume 13, Issue 1, 2017, Pages 57-65, ISSN 2210-8327, https://doi.org/10.1016/j.aci.2016.03.001
24. Jouini M., Rabai L., Aissa A., "*Classification of security threats in information systems*", 5th International Conference on Ambient Systems, Networks and Technologies (ANT-2014), Procedia Computer Science 32 (2014) 489 – 496, ScienceDirect, Elsevier
25. Kandukuri B. R., Paturi R. V. and Rakshit A., (2019), "*Cloud Security Issues*", IEEE International Conference on Services Computing, Bangalore, India, September 21-25, 2019. In Proceedings of IEEE SCC' 2009. pp. 517-520, 2009. ISBN: 978-0-7695-3811-2.
26. Kandukuri, B. R. , Paturi, V. R. and Rakshit A., "*Cloud Security Issues in Cloud Computing: A conceptual analysis and review*" Proceedings of the 2009 IEEE International Conference on Services Computing, Bangalore, India, September 21-25, 2019. In Proceedings of IEEE SCC' 2009. pp. 517-520, 2009. ISBN: 978-0-7695-3811-2. Washington DC, http://dx.doi.org/10.1109/SCC.2009.84
27. Khan, S., et al., 2016. "*Cloud log forensics: foundations, state of the art, and future directions*". ACM Comput. Surv. (CSUR) 49 (1), 7.
28. Kourai, K., Azumi, T., Chiba, S., 2017. "*A self-protection mechanism against stepping-stone attacks, for IaaS clouds*", In: Proceedings of the Ubiquitous Intelligence & Computing and 9th International Conference on Autonomic & Trusted Com- puting (UIC/ATC), 2017 9th International Conference on. 2017. IEEE.
29. Li, T., et al., 2019. "*LARX: Large-scale anti-phishing by retrospective data-exploring based on a cloud computing platform*". In: Proceedings of the Computer Communications and Networks (ICCCN), 2019 Proceedings of 20th International Conference on, 2019. IEEE
30. McMillan, R., 2019. *"Hackers Find a Home in Amazon EC2's Cloud*". URL : ⟨http://www.computerworld.com/article/2521744/security0/hackers-find-a-home-in-amazon-s-ec2-cloud.html⟩.
31. Mell, P. and Grance, T. (2021), "*The NIST Definition of Cloud Computing*", Special Publication (NIST SP), National Institute of Standards and Technology, Gaithersburg, MD, [online], https://doi.org/10.6028/NIST.SP.800-145
32. Minhaj Ahmad Khan (2016), "*A Survey on Security Issues for Cloud Computing", Journal of Network and Computer Applications*", Volume 71, August 2016, Pages 11-29, https://doi.org/10.1016/j.jnca.2016.05.010, Elsevier
33. Modi, C.N., Patel, D.R., Borisaniya, B., Patel, H., Patel, A., & Rajarajan, M. (2018). "*A survey of intrusion detection techniques in Cloud*". J. Netw. Comput. Appl., 36, 42-57.

34. Nasridinov, A., Byun, J.-Y., Park, Y.-H., 2022. “*UNWRAP: An approach on wrapping-attack tolerant SOAP messages*”. In: Proceedings of the Cloud and Green Computing (CGC), 2022 Second International Conference on. 2022. IEEE.
35. Nikolai, J., Wang, Y., 2014. “*Hypervisor-based cloud intrusion detection system*”. In: Proceedings of the Computing, Networking and Communications (ICNC), 2014 International Conference on. 2014. IEEE.
36. Oktay, U., Sahingoz, O., 2015. “*Proxy Network Intrusion Detection System for cloud computing*”. In: Proceedings of the Technological Advances in Electrical, Electronics and Computer Engineering (TAEECE), 2013 International Conference on. 2015. IEEE.
37. Qaisar, S., Khawaja, K.F., 2022. “*Cloud computing: network/security threats and countermeasures*”. Interdiscip. J. Contemp. Res. Bus. 3 (9), 1323–1329
38. Parhizkari, S. (2024). “*Anomaly Detection in Intrusion Detection Systems*”. IntechOpen. doi: 10.5772/intechopen.112733
39. Pearson, S., (2017, May), “*Taking account of privacy when designing cloud computing services*”. In: Proceedings of the 2009 ICSE workshop on software engineering challenges of cloud computing. IEEE computer society, pp 44–52
40. Rodero-Merino,L., et al.,2023., “*Building safe PaaS clouds: a survey on security in multitenant software platforms*”, Comput. Secur. 31(1), 96–108.
41. Reuben, J.S., 2017. “*A Survey on Virtual Machine Security*”. Helsinki University of Technology.
42. Riquet, D., Grimaud, G., Hauspie, M., 2022. “*Large-scale coordinated attacks: impact on the cloud security*”. In: 2022 Sixth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), pp. 558–563. http://dx.doi.org/10.1109/IMIS.2012.76.
43. Rodero-Merino, L., et al., 2022. “*Building safe PaaS clouds: a survey on security in multitenant software platforms*”. Comput. Secur. 31(1), 96–108.
44. Sabahi, F., 2011. “*Virtualization-level security in cloud computing. In: Proceedings of the Communication Software and Networks (ICCSN)*”, 2011 IEEE 3rd International Conference on. 2011. IEEE.
45. Safiia, M., “*An introduction to digital crimes*”, International Journal in Foundations of Computer Science & Technology (IJFCST), Vol.5, No.3, May 2015
46. Sajid, M., Malik, K.R., Almogren, A. et al. “*Enhancing intrusion detection: a hybrid machine and deep learning approach*”. J Cloud Comp 13, 123 (2024). https://doi.org/10.1186/s13677-024-00685-x
47. Scarfone, K., 2021, “*Guide to Security For Full Virtualization Technologies*”, DIANE Publishing.
48. Scarfone, K.A., Mell, P.M., 2017. Sp 800-94. “*Guide to Intrusion Detection and Prevention Systems (IDPS)*”. Technical Report, Gaithersburg, MD, United States
49. Shenfield, A., Day, D., Ayesh, A., 2018, “*Intelligent intrusion detection systems using artificial neural networks*”, ICT Express Volume 4, Issue 2, June 2018, Pages 95-99. https://doi.org/10.1016/j.icte.2018.04.003.
50. Schmidt, F., F. Suri-Payer, A. Gulenko, M. Wallschläger, A. Acker and O. Kao, "*Unsupervised Anomaly Event Detection for Cloud Monitoring Using Online Arima*," 2018 IEEE/ACM International Conference on Utility and Cloud Computing Companion (UCC Companion), Zurich, Switzerland, 2018, pp. 71-76, doi: 10.1109/UCC-Companion.2018.00037.
51. Szefer, J., et al., “*Eliminating the hypervisor attack surface for a more secure cloud*”, In: Proceedings of the Proceedings of the 18th ACM conference on Computer and communications security. 2021. ACM

52. Subashini S, Kavitha V (2021) “A survey on security issues in service delivery models of cloud computing”. J Netw Comput Appl 34(1):1–11
53. Sun, Y., He., D., 2022. “*Model checking for the defense against cross-site scripting attacks*”. In: Proceedings of the Computer Science & Service System (CSSS), 2012 International Conference on 2022. IEEE
54. Tabrizch, H., Kuchaki-Rafsanjani, M., “*A survey on security challenges in cloud computing: issues, threats, and solutions*”, The Journal of Supercomputing (2020) 76:9493–9532, https://doi.org/10.1007/s11227-020-03213-1, Springer.
55. Tandon, S., SB, S., Agrawal, V., 2014. “*Cache-based side-channel attack on aes in cloud computing environment*”. Int. J. Eng. Res. Technol. 3 (10), 1080–1084.
56. Tavallaee, M., Bagheri E., Lu, W., and Ghorbani A. A., "*A detailed analysis of the KDD CUP 99 data set*" 2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications, Ottawa, ON, Canada, 2019, pp. 1-6, doi: 10.1109/CISDA.2009.5356528.
57. Turnbull, L., Shropshire, J., 2023. “*Breakpoints: an analysis of potential hypervisor attack vectors*”. In: Proceedings of the Southeast on IEEE. 2013. IEEE.
58. Wang, Z., Jiang, X., 2020. “*Hypersafe: A lightweight approach to provide lifetime hypervisor control-flow integrity*”. In: Proceedings of the Security and Privacy (SP), 2010 IEEE Symposium on. 2020. IEEE.
59. Wu, H., Ding, Y., Winer, C., Yao, L., 2020. “*Network security for virtual machine in cloud computing*”. In: 2020 5th International Conference on Computer Sciences and Convergence Information Technology (ICCIT), pp. 18–21. http://dx.doi.org/ 10.1109/ICCIT.2010.5711022
60. Xiao, Z., Xiao, Y., 2023. “*Security and privacy in cloud computing*”. Commun. Surv. Tutor. IEEE 15 (2), 843–859.
61. You, P., et al., 2022, “*Security issues and solutions in cloud computing*”. In: Proceedings of the Distributed Computing Systems Workshops (ICDCSW), 2022 32nd International Conference on. 2022. IEEE.
62. X. Zhang, L. Liu, X. Chen, S. Xie and L. Lei, "*A Novel Multitemporal Cloud and Cloud Shadow Detection Method Using the Integrated Cloud Z-Scores Model*," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 12, no. 1, pp. 123-134, Jan. 2019, doi: 10.1109/JSTARS.2018.2889150
63. Zhou, F., Goel, M., Desnoyers, P., Sundaram, R., 2021. “*Scheduler vulnerabilities and coordinated attacks in cloud computing*”. In: 2021 10th IEEE International Symposium on Network Computing and Applications (NCA), pp. 123–130.
64. Zhou M., Zhang R., Xie W., Qian W. and Zhou A., "*Security and Privacy in Cloud Computing: A Survey*" 2010 Sixth International Conference on Semantics, Knowledge and Grids, Beijing, China, 2020, pp. 105-112, doi: 10.1109/SKG.2010.19.

This page was intentionally left as blank