Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет Інфокомунікацій
(повна назва)
Кафедра Інфокомунікаційної інженерії імені В.В. Поповського
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

рівень вищої освіти другий (магістр)

Аналіз методів криптозахисту даних в хмарних обчисленнях
(Analysis of Security Techniques for Protecting Data in Cloud Computing)
(тема)

Виконав:
студент 5 курсу, групи АМСЗІімв-20-1
Якдува Усман Хассан
(прізвище, ініціали)
Спеціальність: 125 Кібербезпека
(код і повна назва спеціальності)
Освітня програма: Адміністративний менеджмент у сфері захисту інформації
of(повна назва освітньої програми)
Керівник: ст.викл. кафедри ІКІ ім. В.В. Поповського
Марчук А.В.
(посада, прізвище, ініціали)

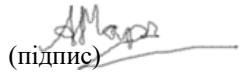
Допускається до захисту
Зав. кафедри _____ Лемешко О.В.
(підпис) (прізвище, ініціали)

2023 р.

Кваліфікаційна робота не містить відомостей, що заборонені до відкритого друку

Студент 5 курсу
групи АМСЗІімв-20-1  Якдува Усман Хассан

(підпис) (ініціали, прізвище)

Керівник  А.В. Марчук

(підпис) (ініціали, прізвище)

Харківський національний університет радіоелектроніки

Факультет Інфокомунікацій
(повна назва)
Кафедра Інфокомунікаційної інженерії імені В.В. Поповського
(повна назва)
Рівень вищої освіти другий (магістр)
Спеціальність 125 Кібербезпека
(код і повна назва)
Освітня програма Адміністративний менеджмент у сфері захисту інформації

(повна назва)

ЗАТВЕРДЖУЮ
Зав. кафедри_____
(підпис)
«____» _____2023 р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

студенту Якдува Усман Хассан
(прізвище, ім'я, по батькові)

1. Тема роботи: Аналіз методів криптозахисту даних в хмарних обчисленнях (Analysis of Security Techniques for Protecting Data in Cloud Computing)
затверджена наказом по університету від «_31__» __10__ 2022 р. №_1434_ Ст.
2. Термін подання студентом роботи до екзаменаційної комісії _22.01. 2023 р.
3. Вихідні дані до роботи: Вимоги стандарту ISO / IEC 27001 A.10.10.2 , щодо виявлення атак. Допустимість використання програмного та апаратного забезпечення (стандарт ISO / IEC 27001 A.12.4.1), контроль змін в інформаційній системі (ISO / IEC 27001 A.12.5.1),), а також основні протоколи, що забезпечують функціонування мереж IoT з використанням хмарних технологій.
4. Перелік питань, що потрібно опрацювати в роботі:
1) Огляд поточного стану архітектури, технологій і протоколів, які забезпечують функціонування хмарних обчислень.
2) Аналіз вразливостей систем хмарних обчислень.
3) Дослідження безпеки систем обробки конфіденційних данних у хмарних обчисленнях.

5. Перелік графічного матеріалу із зазначенням креслень, плакатів, комп'ютерних ілюстрацій: Демонстраційний матеріал у вигляді ppt-презентації

6. Консультанти розділів роботи

Найменування розділу	Консультант (посада, прізвище, ім'я, по батькові)	Позначка консультанта про виконання розділу	
		(підпис)	(дата)
Основна частина	ст.викл. Марчук Артем Володимирович		22.01.23

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1	Отримання завдання	31.10.22	Виконано
2	Збір матеріалів для дослідження	15.11.22	Виконано
3	Розробка 1 розділу	20.12.22	Виконано
4	Розробка 2 розділу	01.01.23	Виконано
5	Розробка 3 розділу	15.01.23	Виконано
6	Оформлення кваліфікаційної роботи	22.01.23	Виконано

Дата видачі завдання 31 жовтня 2022 року

Студент _____ Якдува Усман Хассан
(підпис) (прізвище, ініціали)

Керівник роботи _____ ст.викл. Марчук А.В.
(підпис) (посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка: 53 с., 15 рис., 3 табл., 12 джерел.

ІНФОРМАЦІЙНА БЕЗПЕКА, ХМАРНІ ОБЧИСЛЕННЯ, ХМАРНЕ ЗБЕРІГАННЯ, ЗАГРОЗИ ІНФОРМАЦІЙНІЙ БЕЗПЕЦІ, ХМАРНІ ВРАЗЛИВОСТІ, ПІДСИСТЕМА АУДИТУ ДАНИХ, АЛГОРИТМ ЗАХИСТУ ДАНИХ, ДОСЛІДЖЕННЯ ЕФЕКТИВНОСТІ.

Об’єкт дослідження – процеси функціонування телекомунікаційних систем з використанням хмарних обчислень.

Предмет дослідження – характеристики телекомунікаційних систем з використанням хмарних обчислень.

Мета роботи – аналіз архітектури і параметрів технологій телекомунікаційних систем з використанням хмарних обчислень, а також вдосконалення методів інформаційного захисту таких систем.

Методи дослідження – емпіричний аналіз, порівняння, аналітичне моделювання.

На сьогоднішній день тема захисту даних в мережах з використанням хмарних обчислень є актуальною, оскільки в наш час відбувається суттєве збільшення кількості користувачів і потоків інформації в системах хмарних обчислень.

У роботі виконано аналіз сучасних систем хмарних обчислень, їх загроз та вразливостей. Запропоновано удосконалення методу інформаційного захисту даних в хмарних обчисленнях. Виконано дослідження продуктивності алгоритму захисту даних у хмарному сховищі. Розроблені рекомендації по збільшенню захисту систем хмарних обчислень.

ABSTRACT

Explanatory note: 53 p., 15 fig., 3 tables, 12 sources.

INFORMATION SECURITY, CLOUD COMPUTING, CLOUD STORAGE, INFORMATION SECURITY THREATS, CLOUD VULNERABILITIES, DATA AUDIT SUBSYSTEM, DATA PROTECTION ALGORITHM, PERFORMANCE STUDY.

The object of research is the processes of functioning of telecommunication systems using cloud computing.

The subject of the research is the characteristics of telecommunication systems using cloud computing.

The purpose of the work is to analyze the architecture and technology parameters of telecommunication systems using cloud computing, as well as to improve the methods of information protection of such systems.

Research methods – empirical analysis, comparison, analytical modeling.

Today, the topic of data protection in cloud computing networks is relevant, as nowadays there is a significant increase in the number of users and information flows in cloud computing systems.

The paper analyzes modern cloud computing systems, their threats and vulnerabilities. It is proposed to improve the method of information protection of data in cloud computing. A performance study of the data protection algorithm in cloud storage was performed. Recommendations for increasing the protection of cloud computing systems have been developed.

CONTENTS

LIST OF ABBREVIATIONS, SYMBOLS, UNITS AND TERMS

ABAC – Attribute-Based Access Control
ACL – Access Control List
AES – Advanced Encryption Standard
API – Application Programming Interface
AWS – Amazon Web Services
CRM – Customer Relationship Management
CSA – Cloud Security Alliance
CSP – Cloud Service Provider
DBaaS – Database as a Service
DC – Data Center
DoS – Denial of Service
DDoS – Distributed Denial of Service
ERP – Enterprise Resource Planning
GCP – Google Cloud Platform
HaaS – Hardware as a Service
IaaS – Infrastructure as a Service
IP – Internet Protocol
IT – Information Technology
NIST – National Institute of Standards and Technology
PaaS – Platform as a Service
PN – Personal Network
PTA – Practical Threat Analysis
QoS – Quality of Service
RSA – Rivest, Shamir and Adleman
SaaS – Software as a Service
SHA – Secure Hash Algorithm
TAM – Threat Analysis and Modeling
TCP – transmission control protocol
TPA – Third-Party Auditor
VM – Virtual Machine
VPC – Virtual Private Cloud

VPN – Virtual Private Network

UDP – user datagram protocol

INTRODUCTION

From the past few years, there has been a rapid progress in Cloud Computing. Cloud Computing delivers a great variety of resources like computational power, storage, computational platforms, and applications to users via the Internet. The major Cloud providers in the current market segment are Amazon, Google, IBM, Microsoft, Salesforce, etc.

The use of cloud computing significantly reduces the costs of companies and gives them the ability to flexibly respond to changing the computing needs. You can work on software provided by the cloud. You can store large amounts of data. You don't need expensive hardware to store your databases. It is possible to use and run various applications and online services. Users can perform complex calculations and perform simulations of complex systems. It is very convenient to work remotely. In this case, users can use inexpensive hardware and pay as they grow. In such laptops and desktop computers, it is not necessary to have large data storages, high-speed processors and large RAM.

Along with the advantages of cloud services, there are also a number of disadvantages. While the cost of renting cloud services is still quite high. However, every year the cost of rent is gradually decreasing. Competition between cloud service providers contributes to this process. There is a lack of user awareness of the opportunities provided by cloud service providers. These shortcomings hinder the development of cloud services. In addition, the use of cloud services requires access to the Internet. However, despite the presence of shortcomings, this technology has a great future.

In connection with the increase in the number of users of cloud services, the task of ensuring the protection of data stored in cloud storages becomes relevant. On the other hand, it is important to provide secure access to cloud services. User authentication, issues of information encryption and decryption, protection of Internet access channels.

In this Master Thesis, we provided a general overview of cloud computing, its types, delivery models and features. Also we analyzed threats and vulnerabilities of cloud computing systems and proposed improvements of security techniques for protecting data in them.

It has been proposed to combine the cloud storage data protection algorithm with a third-party public audit scheme and store encryption keys separately from the cloud storage. For this improved algorithm we made a program implementation and investigated its performance to prove that it can be used in real cloud data storage.

1 ANALYSIS OF MODERN CLOUD COMPUTING

1.1 General overview

Cloud Computing term (also the term "cloud (distributed) data processing") is usually used to mean the provision of computer resources and capacities to the user in the form of an Internet service. Thus, computing resources are provided to the user in a "clean" form, and the user may not know which computers handle its requests, under which operating system it is happening etc. By cloud technologies we will understand the model of system procedures that ensure a targeted change of material and virtual objects through the network which can be promptly provided and released at the request of the user with minimal operational costs and / or calls to the provider.

Clouds are often compared to mainframes, finding a lot in common between them. The fundamental difference between the cloud and mainframes is that its computing power is theoretically unlimited. The second fundamental difference is that, simply speaking, terminals for mainframes are used only for user interaction with a task launched for processing. In the cloud, the terminal itself is a powerful computing device capable not only of accumulating intermediate information, but also of directly managing the global system of computing resources.

Among the earlier (in the 1990s) data processing technologies, the so-called grid computing has become somewhat widespread. This direction was initially considered as an opportunity to use free resources of processors and to develop a system of voluntary rental of computing capacities. A number of projects (GIMPS, distributed.net, SETI@home) have proven that this model of calculations is quite effective. This technology is used to solve scientific and mathematical problems that require significant computing resources. It is known that grid computing is also used for commercial purposes. For example, with their help, some time-consuming tasks related to economic forecasting, analysis of seismic data, development and study of the properties of vaccines and new drugs are performed. Indeed, grid computing and clouds have many similar features in architecture and applied principles. However, the cloud computing model is considered more promising today due to a much more flexible platform for working with remote computing resources.

Nowadays, large computing clouds consist of thousands of servers located in data centers (DC). They provide resources for tens of thousands of applications that are used simultaneously by millions of users [1]. Cloud technologies are a convenient tool for enterprises that are too expensive to maintain their own ERP, CRM or other servers that require the purchase and configuration of additional equipment.

Due to their convenience, such cloud services as, for example, those provided by Google ("Documents", "Calendar", etc.) are gradually becoming widely used among private users.

The reasons for the growing popularity of cloud technologies are clear: the possibilities of their application are very diverse and allow saving both on maintenance and personnel, and on infrastructure. Hardware can be greatly simplified when processing data and storing information in remote data centers. All these problems are almost entirely transferred to the service provider. In addition, this approach allows you to standardize software, even if different operating systems (Windows, Linux, MacOS, etc.) are installed on the company's computers. Cloud technologies make it easier to provide access to company data for both customers and employees who are outside the office but have the ability to connect via the Internet.

It is clear that using cloud computing is much more convenient. The most important disadvantage that can be immediately noticed is the complete dependence on the provider of these services. In fact, the enterprise (user) turns out to be a hostage of the service provider and the Internet access provider. Although the reliability of cloud computing providers is increasing, it is necessary to make a lot of effort to ensure the reliability and security of data, for example, to have redundant communication channels, duplicating capacities for the possibility of switching to them and, of course, to think about the availability of information and security. In addition, cloud computing is absolutely not suitable for enterprises related to state and military secrets. No commission will issue a certificate for such a system when working with information that is not subject to disclosure.

1.2 Main features of cloud technologies

NIST (National Institute of Standards and Technology, USA) in its document "The NIST Definition of Cloud Computing" defines the following characteristics of clouds:

First, "broadband network access." Access to resources in the cloud can be obtained using several types of devices. This includes not only the most common devices (laptops, workstations, etc.), but also mobile phones, thin clients, and so on. Contrast "broadband network access" with access to computing and network resources in the mainframe era. Computing resources forty years ago were scarce and expensive. Their use is limited due to workload priorities and importance in order to conserve these resources. Similarly, network resources are also limited. Internet Protocol (IP)-based networks were widespread at the time, so high-bandwidth, low-latency networks did not exist. Over time, costs related to the network (eg, costs related to computing and data storage) have decreased due to the scale of production and commercialization of the relevant technologies. As the bandwidth of the network increased, access to it and its scalability increased accordingly. "broadband network access" can be considered both a characteristic of cloud computing and a contributing factor to its development.

"Self-service on demand" is a key - some say essential - characteristic of cloud computing. Viewing IT as a complex supply chain with the application and the end user at the end of the chain, the ability to self-sustain resources in typical IT environments fundamentally disrupts the workflow and processes that have evolved as a function of enterprise IT over the past few decades. This includes workflows related to the purchase of storage systems, servers, network nodes, software licenses, and so on. Self-provisioning in non-cloud or legacy architectures forces traditional processes and functions such as bandwidth planning, network management (quality of service assurance or QoS) and security (creation of access control lists or ACLs) to stop or even break down completely. The well-known effect in supply chain management - when incomplete or inaccurate information leads to high variability of production costs - applies not only to the production environment, but also directly to the allocation of IT resources in a non-cloud environment (2). Cloud architectures, however, are designed and created with self-sufficiency in mind. This premise implies the use of fairly complex software frameworks or portals to manage the initialization and back office functions. Historically, the lack of commercial off-the-shelf (COTS) software specifically designed to automate cloud computing has forced many companies to create their own portals and frameworks to support these functions. COTS software packages designed for corporate management and automated enterprise workloads currently make up a significant portion of the overall potential cloud computing market.

Pooling of resources is a fundamental prerequisite for scalability in the cloud. Without converged computing networks and storage, a service provider must deliver services over multiple isolated or discrete, independent resources with little or no connectivity. The basis of public cloud infrastructures are multi-user environments in which multiple customers share adjacent resources in the cloud with their peers. With multi-tenancy, there is an inevitable increase in operating costs, which can be reduced by certain hardware configurations and software solutions, such as application and server profiles.

"Measured service" means that the use of these "pooled resources" is monitored and reported to the consumer, ensuring transparency of consumption rates and costs. Monitoring usage for chargeback purposes (or simply for cross-departmental reporting and budgeting) has long been a requirement for IT stakeholders, but building such a system is typically not the core competency of most IT departments.

The final characteristic highlighted in NIST's definition of cloud computing is "rapid elasticity." Elastic computing is critical to reducing costs and time to market (TTM). Indeed, the concept of flexible resources in the IT supply chain is so desirable that Amazon has named its cloud platform "Elastic Compute Cloud" ("EC2"). In terms of FTE costs, operational costs associated with resource allocation (movement, addition, change, or MAC) in the IT supply chain typically account for the largest portion of application deployment costs.

1.3 Deployment models of cloud computing

According to the deployment model, clouds are divided into private, public and hybrid (fig 1.1).

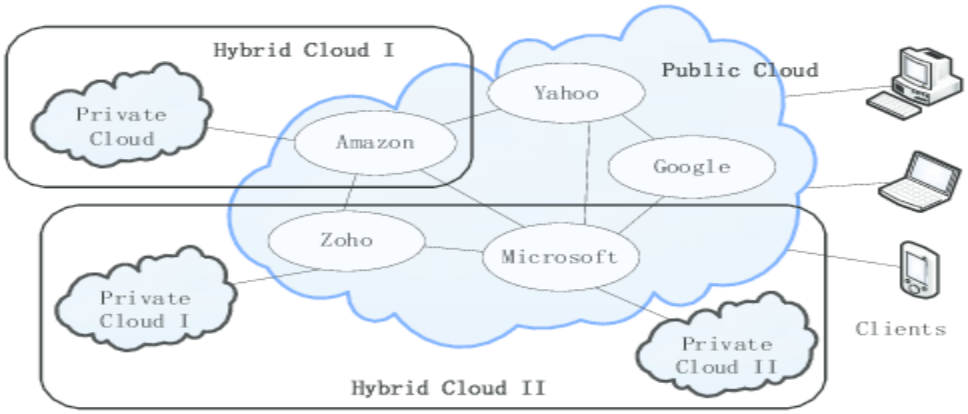


Figure 1.1 - Private, public and hybrid clouds

Private clouds - are the company's internal cloud infrastructure and services. These clouds are located within the corporate network. An organization can manage the private cloud itself or entrust this task to an external contractor. The infrastructure can be located either on the customer's premises, or at an external operator's, or partly at the customer's and partly at the operator's. The ideal version of a private cloud is a cloud that is deployed on the territory of the organization, served and controlled by its employees. Private clouds have the same privileges as public clouds, but with one important feature: the company itself is engaged in the installation and maintenance of the cloud. The complexity and cost of creating an internal cloud can be very high, and the cost of operating it can exceed the cost of using public clouds.

It should be noted that private clouds have advantages over public clouds: more detailed control over various cloud resources provides the company with any available configuration options. In addition, private clouds are ideal when you need to perform work that cannot be entrusted to the public cloud for security reasons.

Public clouds - are cloud services provided by a provider. They are outside the corporate network.

Data clouds users do not have the ability to manage the data cloud or maintain it, all responsibility rests with the owner of this cloud. A cloud service provider assumes responsibility for the installation, management, provisioning, and maintenance of software, application infrastructure, or physical infrastructure. Customers pay only for the resources they use. Any company and individual user can become a subscriber of the offered services. They offer an easy and affordable way to deploy websites or business systems with great scalability that would not be available with other solutions. Examples: online services Amazon EC2 and Amazon Simple Storage Service (S3), Google Apps / Docs, Salesforce.com, Microsoft Office Web. At the same time, public cloud services are mainly provided in the form of standard configurations, that is, based on the conditions of the most common use cases. This means that the user has fewer opportunities to choose the configuration compared to systems in which the resources are managed by the consumer himself. It should also be noticed that because consumers have little control over the infrastructure, processes that require strict security measures and compliance with regulatory requirements are not always suitable for implementation in the public cloud.

Hybrid clouds are a combination of public and private clouds. They are usually created by the enterprise, and the responsibility for managing them is shared between

the enterprise and the public cloud provider. A hybrid cloud provides services, some of which are publicly available, and some of which are private. Typically, this type of cloud is used when an organization has seasonal periods of activity. In other words, as soon as the internal IT infrastructure cannot fulfill current tasks, part of the capacity is transferred to the public cloud (for example, large volumes of statistical information, which in their raw form do not represent value for the enterprise), as well as to provide users with access to enterprise resources (to a private cloud) through a public cloud.

A well-designed hybrid cloud can serve security-critical processes, such as receiving payments from customers, as well as more secondary ones. The main disadvantage of this type of cloud is the difficulty of effectively creating such solutions and managing them. It is necessary to receive services from various sources and organize them as if they were a single source. interactions between private and public components can further complicate the decision. As this is a relatively new architectural concept in the field of cloud computing, new best practices and tools are emerging for this model, and its widespread use may be delayed until it is better studied. According to Tom Bittman, vice president and leading analyst of the American research and consulting company "Gartner", among the three cloud deployment models listed above, private clouds are the most relevant for business at the moment. Bittman singled out five main points that help to get a more accurate idea of the structure of a private cloud.

Depending on the type of cloud service, it can be owned and managed by both the provider and the user, or both. Access rights to resources may also differ (See Table 1).

Table 1.1 - Maintenance and management of various types of cloud resources

Cloud type	Who maintain infrastructure	Owner of the infrastructure	Where the infrastructure placed	Who have access
Public	External provider	External provider	External provider	User
Private	User or External provider	User or External provider	External provider or User	Authorized user
Hybrid	User and External provider	User and External provider	External provider and User	Authorized user and other external users

The cloud is not only virtualization. Although virtualization of servers and infrastructure is an important foundation of private cloud computing, virtualization and management of the virtualized environment itself is not yet a private cloud. Virtualization allows better structuring, pooling and dynamically providing infrastructure resources: servers, desktops, storage containers, network equipment, connecting software, etc. But in order for the environment to technically be considered cloud, other components are needed, such as virtual machines, operating systems or containers of connecting software, highly stable operating systems, grid computing software, software for abstracting storage resources, scaling and clustering tools.

The "private cloud" term, as opposed to public or hybrid, refers to resources used by a single organization, or means that an organization's cloud resources are completely isolated in the cloud from others. The cloud is not necessarily a source of savings. One of the biggest misconceptions is that the cloud will save money. Savings are possible, but not a mandatory attribute. A private cloud allows you to more efficiently reallocate resources to meet corporate requirements and can reduce capital equipment costs. But a private cloud requires an investment in automation, and the savings alone may not cover the entire cost. So, cost reduction is not the main advantage of this model. From this point of view, the main incentive for the implementation of the cloud model should not be savings, but the speed of market entry, the possibility of rapid adaptation and dynamic scaling according to demand, which allow to increase the speed of introduction of new services. The private cloud is not always implemented by the customer. A private cloud means privacy, not a specific location, resource ownership, or self-management. Many providers offer non-local private clouds, that is, allocate resources to a single customer, eliminating the sharing of the same pool by several customers. "The cloud is called private because of its privacy, not because of where it is deployed, who owns it and is responsible for managing it," Bittman emphasizes. Some, for example, can mix their data centers with hosting providers or pool the resources of different customers, but isolate them from each other using a virtual private network (Virtual Private Network - VPN) and other similar technologies. Private cloud (as well as public cloud) is not only infrastructure services. Server virtualization is a big trend and therefore a powerful engine of private cloud computing. But private cloud is not just about infrastructure as a service (IaaS). For example, for developing and testing new software, PaaS makes more sense than simply providing virtual machines.

Today, the fastest growing segment of cloud computing is IaaS. It provides the most low-level data center resources in an easy-to-use form, but does not fundamentally change the principles of operation. To create new cloud-first applications that provide completely new services that may be very different from what previous applications provided, developers are more comfortable using PaaS. A private cloud can stop being private. On the one hand, a private cloud offers the advantages of the cloud: speed of reconstruction, scalability and efficiency, and eliminates some of the security threats, potential and real, that are characteristic of public clouds. On the other hand, over time, the level of service, security and compliance control in public cloud services will definitely increase. Therefore, some private clouds may well move into the category of public clouds. The majority of private cloud services will most likely evolve into hybrid cloud services, expanding available opportunities through the use of publicly available cloud services and other third-party resources.

1.4 Service models of cloud computing

Currently, it is customary to distinguish three main service models of cloud technologies, which are sometimes called cloud layers. We can say that these three layers (Fig 1.2) - infrastructure services, platform services and application services - reflect the structure of not only cloud technologies, but also information technologies in general. Let's dwell on each of them in more detail.

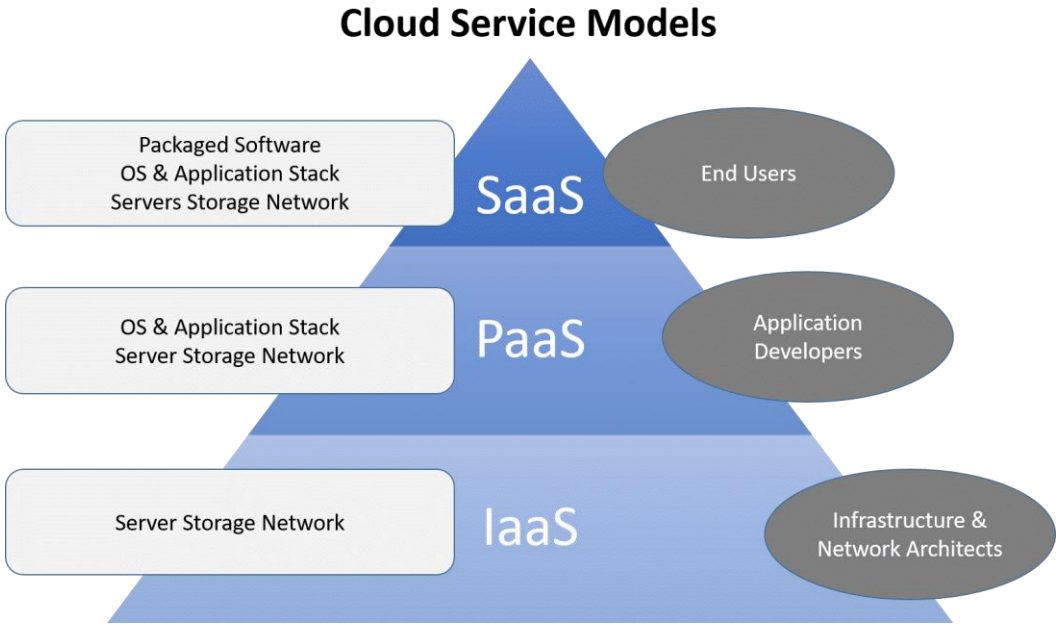


Figure 1.2 - Cloud service models as layers

Infrastructure as a Service - (IaaS) is a set of physical resources, such as servers, network equipment and storage, offered to the customers as services that they request for. Infrastructure services solve the problem of properly equipping the data center, providing computing power as needed. Typically, these services support infrastructure and a much larger number of consumers compared to application services. A private example of infrastructure services is hardware as a service (HaaS). As a service, the user receives equipment on the basis of which he deploys his own infrastructure using the most suitable software.

The consumer does not manage the underlying cloud infrastructure, but has control over operating systems, storage systems, deployed applications, and possibly limited control over the selection of network components (eg, a host with network screens). In this case, the protection of platforms and applications is provided by the consumer himself, and the cloud provider must organize the protection of the infrastructure. Virtualization is often used to provide resources on demand. Benefits. Reduction of capital investments in hardware. Because virtualization techniques are commonly used in this model, savings can be achieved through more efficient use of resources. Reducing the risk of investment loss and implementation threshold, the possibility of smooth automatic scaling. Disadvantages: Business efficiency and productivity are highly dependent on supplier capabilities. Potentially large long-term costs are likely to be required.

Centralization requires new approaches to security measures. Examples of infrastructure services include IBM SmartCloud Enterprise, VMWare, Amazon EC2, Windows Azure, Google Cloud Storage, Parallels Cloud Server, and many others.

Platform as a Service - (PaaS) is a service model in which applications (created or purchased) are provided to the consumer as a set of services. This includes, in particular, middleware as a service, messaging as a service, integration as a service, information as a service, communication as a service, etc. For example, Workplace as a Service (WaaS) allows a company to use cloud computing to organize the workplaces of its employees by configuring and installing all the necessary software for staff work. Data as a Service (DaaS) provides the user with disk space that he can use to store large amounts of information. Security as a Service (SaaS) enables users to quickly deploy products that allow secure use of web technologies, email security, and local system security. This service allows users to save on the deployment and maintenance of their own security system. In other words, the PaaS model is IaaS together with the operating

system and its application programming interface (API - Application Programming Interface). At the same time, the consumer does not manage the basic infrastructure of the cloud, including networks, servers, operating systems and data storage systems, but has control over the deployed applications and possibly some configuration parameters of the hosting environment. Thus, the consumer must take care to ensure the protection of the applications that will be deployed on the provided platforms[3].

Applications can work both in the cloud and in traditional enterprise data centers. To achieve the scalability required in the cloud, the various services offered are often virtualized, just like the infrastructure services discussed earlier. Benefits. Smooth rollout of versions. Smoothness means that, ideally, the user should have little or no experience of software changes in the cloud. Disadvantages As with the previous service model, centralization requires robust security measures. Examples of platform services are IBM SmartCloud Application Services, Amazon Web Services, Windows Azure, Boomi, Cast Iron, Google App Engine, and others.

Software as a Service - (SaaS) assumes access to applications as a service, that is, the provider's applications are launched in the cloud and provided to users on demand as services. In other words, the user can access the software deployed on remote servers using the Internet, and all issues of updating and licenses for this software are regulated by the provider of this service. Payment in this case is made for the actual use of the software. Sometimes these services are made free by providers, as they have the opportunity to generate revenue, for example, from advertising. The application is accessible through various client devices or through thin client interfaces, such as web browsers, or webmail, or application interfaces. At the same time, the consumer does not manage the basic infrastructure of the cloud, including networks, servers, and operating systems. Ultimately, you are solely responsible for saving access parameters (logins, passwords, etc.) and following the provider's recommendations for secure application settings. Application services are most familiar to the everyday user. The most common examples of applications of this type are mail services GMail, Yahoo Mail. In total, there are thousands of SaaS applications, and thanks to Web 2.0 technology, the number is growing every day. Among the application services, there are many applications aimed at the enterprise community. There is software that manages payroll, human resources, teamwork, relationships with clients and business partners, etc.

Advantages: Reduction of capital investments in hardware and labor resources; reducing the risk of investment loss; smooth iterative update.

Disadvantages: As with the previous two models, centralization requires robust security measures. Examples of SaaS include Gmail, Google Docs, Netflix, Photoshop.com, Acrobat.com, Intuit QuickBooks Online, IBM LotusLive, Unyte, Salesforce.com, Sugar CRM, and WebEx. A large part of the growing mobile application market is also a SaaS implementation.

Service models by means of content access can be seen in Table 1.2.

Table 1.2 - Service models by ways of access and management

Service models	Access and control ways	Contents
SaaS	Web browser	Cloud applications: social networks, office applications, content management systems, intelligent data processing.
PaaS	Cloud development environment	Cloud platform: programming languages, libraries, service composition configuration utilities, structured data.
IaaS	Management system of virtual infrastructure	Cloud infrastructure: computing servers, data storage, network connection establishment (firewalls, load balancing)

However, these are not all cloud technology service models. There are also.

1**.** Hardware as a service (HaaS) is a procurement model similar to leasing or licensing, where hardware owned by a managed service provider (MSP) is installed on the customer's premises, and a service level agreement (SLA) defines the responsibilities of both parties . Sometimes the customer pays a monthly fee for the use

of the hardware, sometimes the use is included in the MSP's hardware installation, monitoring and maintenance fee structure. In any case, if the equipment fails or becomes obsolete, the MSP is responsible for decommissioning and replacing it. Depending on the terms of the SLA, decommissioning may include erasing proprietary data, physically destroying hard drives, and confirming that old equipment has been legally recycled.

The HaaS model can be a cost-effective way for small and medium-sized businesses to provide employees with state-of-the-art equipment at minimal costs. HaaS can be contrasted with infrastructure-as-a-service (IIS) and managed hosting procurement models, in which the hardware is hosted on the MSP's premises.

2**.** DBaaS (Database as a Service) is a type of PaaS. Using DBaaS, a user can access any type of database on demand. The user can quickly deploy a database on any class of equipment in the environment of the software platform (operating system) chosen by him. [4]

The user can choose a database, specifying its version, general configuration, and a number of other features (for example, placement). On request, the DB can be placed in the OS on a virtual machine or connected within a container.

Conclusions to chapter 1

In this chapter, we gave an analysis of modern cloud technologies, various models of cloud services, a definition of cloud computing, and considered the architecture that allows cloud services to be as they are.

The relevance of the use of cloud technologies was also revealed, which leads to the question of the existence of certain threats that are associated with the existence of these technologies. Existing threats to cloud services will be discussed in the next section. And in the third section, protection mechanisms will be considered.

2 THREATS AND VULNERABILITIES OF CLOUD COMPUTING

Cloud computing as a new paradigm of distributed computing provides on-demand services, reducing capital investment in infrastructure and maximizing the use of available resources. Cloud technologies provide mobility of applications and infrastructure services, as well as independence of physical / hardware platforms from existing distributed computing and network applications [5]. With the growth of cloud computing and the availability of computing resources for IT consumers, the industry is becoming more flexible, capable and cost-effective for developing and hosting applications. However, when adopting this powerful new system, security was the most important and critical issue. To understand security issues, it is necessary to analyze threats, vulnerabilities, and risks as different factors affecting the security of cloud computing.

Threat modeling in the form of a systematic and comprehensive analysis of threats and vulnerabilities is necessary to ensure the confidentiality, integrity and availability of a cloud computing security deployment. Threat modeling collects basic information needed in the form of usage scenarios, external dependencies, implementation assumptions, details of internal and external security implementation [6]. A number of threat modeling methods have been developed to assess and analyze such threats and vulnerabilities as:

- Microsoft threat modeling is a model of an effective process that includes five logical steps from the classification of assets to the elimination and classification of threats and vulnerabilities.

- Microsoft's Threat Analysis and Modeling (TAM) [6] is a model based on the business objectives to be achieved by the application. TAM tools are used to generate and classify threats by measuring their harmful effects on system components.

- Practical threat analysis (Practical Threat Analysis, PTA) [7] determines an effective risk reduction plan for a specific system architecture to obtain the value of system assets.

- The structure and methodology of the threat model for personal networks (PNs) [8]. This threat model, based on personal network analysis, gives a good idea of the system's assets and their value. Defining all resources using UML diagrams makes it possible to protect data and network functions from any threats.

- Modeling threats in a widespread computing paradigm by developing cloud computing. Each user faces different security domains with multiple identities. The above model is a new threat simulation that includes the problem of a widespread computing environment.

Unfortunately, due to the growth of cloud environments as large distributed computing, all of the above techniques do not include all problems, therefore, to solve the problems of ubiquitous computer security, it is necessary to develop new approaches to threat modeling. Before going into more detail to consider threats, vulnerabilities and solutions specific to cloud environments, we will define vulnerability and threat as follows:

A threat is damage or unauthorized access that could result from a vulnerability and destroy an organization's assets, operations, or system information.

Vulnerability is any weakness of an information system, system security procedures, internal control or implementation that can be exploited or caused by threat resources [9].

2.1 Threats of cloud computing

The Cloud Security Alliance (CSA) [10] as a security service provider has released a security guide for critical areas in cloud computing to manage risks and understand security threats. The most significant threats associated with the on-demand nature of cloud computing are categorized as follows:

- Data loss or leakage (T1): This threat is rated as the most serious and dire threat to businesses and consumers. Any deletion of data by the service provider or an accident, such as a fire, may result in the loss of the consumer's data.

- Account or service theft (T2): This vulnerability allows attackers to steal credentials and access to critical areas of cloud computing services. The organization must prohibit the sharing of credentials between different services and users and use reliable authentication methods.

- Insecure interface (T3): Cloud computing clients use application programming interface (API) or programming interfaces to interact with and manage cloud services. Authentication, access control, and monitoring technologies for APIs protect computing resources from malicious attacks.

- Denial of Service (T4): Distributed Denial of Service (DDoS) is the main threat to availability security when it comes to increasing the reliability of organizations on public cloud services [11]. On the other hand, this attack prevents users from accessing their data or applications and prevents them from reaching their destination. Cloud service providers must be confident in availability protection, and customers - in the level of availability protection within the provider.

- Malicious Insider (T5): The system has been compromised by an authorized employee, business partner, or administrator who has access to the network or resources. This dangerous threat affects the confidentiality, integrity and availability of business information.

- Data leakage (T6): One of the worst situations for every organization is unauthorized access or illegal viewing of data by competitors. Encrypting your data can reduce the risk of this threat, but you should be careful with the encryption key because if you lose it, you lose your data.

- Abuse of cloud services (T7): Cloud computing providers do not have strict registration procedures, and any user with a valid credit card can register for cloud services [12]. Integrating weak registration fraud detection methods allows an attacker to effectively exploit data through aggressive cloud models such as PaaS and IaaS.

- Lack of due diligence (T8): Cost reduction, access to a pool of resources and improved security are the most important factors that can be useful for an organization to accelerate the development of cloud computing. Without understanding the Cloud Service Provider (CSP) environment, mismatched expectations have been established as a critical cloud security control issue. However, in order for resources to be sufficiently skilled, organizations must understand service providers' offerings and risks.

- Insecure migration of virtual machines (T9): By moving different virtual machines during hybrid and converged clouds, attackers can gain illegal access to data and transfer it to an untrusted host. Virtualization as a core component of IaaS is the main target of attackers' attacks. Reliable cloud computing and encryption technology protect data resources from dangerous migration of virtual machines.

The threat of common vulnerabilities exists across all models because the underlying components that deploy infrastructure, platforms, and applications do not provide strong isolation between cloud computing models.

2.2 Vulnerabilities of cloud computing

When transferring critical data and applications of the organization to cloud services, one should take into account their various significant vulnerabilities, the main characteristics of the cloud, known security controls and the most modern cloud offers.

- Session Hijacking (V1): Session Hijacking refers to hackers sending commands to a web application to gain unauthorized access to information, or exploiting weaknesses in a web service to allow a hacker to perform actions such as deleting user data or sending spam to the network via the Internet.

- Virtual machine traversal (V2): This vulnerability allows an attacker to run code on a virtual machine that allows the operating system to compromise and interact directly with the hypervisor to access the host operating system and other virtual machines. For prevention, the system must detect malicious activity at the level of virtual machines.

- Legacy cryptography (V3): The development of insufficiently strong encryption or its absence allows an attacker to decrypt encrypted data. To protect the system from this vulnerability, the user must be sure that the real data is encrypted, use proper key storage, and design a good algorithm.

- Unauthorized access to the management interface (V4): The cloud management interface has access to users of cloud services to manage services on demand. Unauthorized access allows attackers to gain complete control over users and applications.

- Internet Protocol (V5): The lack of authentication methods not included in the basic protocol allows attackers to inject their malicious traffic into the network. On the other hand, IP or related protocols such as UDP and TCP are vulnerable to various types of Denial of Service (DoS) attacks, including session hijacking.

- Data recovery (V6): Cloud computing allows you to allocate or redistribute resources between different users. This elastic characteristic can lead to data theft, data leakage and other security threats. Most organizations use third-party providers for data recovery, so they must consider the security risk when dealing with data with external companies and ensure that the service provider is properly vetted

- Billing (V7): Cloud computing meters and metering services such as storage, user account and processing are used to optimize service delivery. Applicable

vulnerabilities include the processing of accounting and billing data, as well as the leakage of invoices.

- Vendor lock-in (V8): - this is a situation where a cloud user is dependent on one vendor and cannot deal with another vendor without significant inconvenience. The lack of standards is the main reason why users cannot easily switch from one provider to another.

2.3 Comparison of existing threat modeling frameworks

Before considering the proposed model, we illustrate the details of the existing models in Table 2.1 in order to analyze and compare the existing models with the proposed model. Table 2.1 is based on the characteristics of the threat model that will be used in cloud computing.

Table 2.1 - Comparison of existing threat models

Characteristics of the threat model	Threat modeling from Microsoft	TAM	RTA	Threat model for personal networks	Threat Modeling in a Comprehensive Computational Paradigm	Model Amini-Jamil
Identification and classification of assets	●		●	●	●	●
Setting user roles						●
Identification of security areas					●	●
Reliability assessment					●	●
Scanning security areas						●
Threat detection	●	●		●	●	●
Vulnerability detection				●	●	●
Implementation of a countermeasure					●	●

Continuation of table 2.1

Characteristics of the threat model	Threat modeling from Microsoft	TAM	RTA	Threat model for personal networks	Threat Modeling in a Comprehensive Computational Paradigm	Model Amini-Jamil
Threat assessment and measurement	●	●		●	●	●
Assessment and measurement of vulnerabilities				●		●
Identification of new assets, threats, vulnerabilities					●	●

The cloud computing security model by Amini-Jamil is presented in Figure 2.1. This model consists of four main stages, each of which includes several subsections. The first step is to identify the assets and clarify who or what has access to them. Trust, represented by the term trust, is used to refer to a variety of redundant web services to indicate qualified trust between a user and a service provider, or between different providers. Then, at the second stage, the provider's ability to meet user requirements in the literal sense of the word is determined. Identifying unique threats and eliminating them by developing appropriate countermeasures was demonstrated in the third stage. The main goal of this stage is to identify and further identify new threats to improve security. At the last stage, a system rating is presented to identify the most dangerous and effective threats and vulnerabilities.

Define assets. An IT asset is the data, software or hardware owned by the company and used to run the business. Organizations must be confident that these resources are accessed by authorized users and that they are configured to use the latest security technologies to protect against security threats and vulnerabilities. The most unauthorized access or effective changes in the configuration management process are made by an unofficial user to the system. Therefore, in order to block the attack, it is necessary to know the machines of enterprises and their location. Today, organizations

use feasible and effective asset management (SAM) tools to analyze their own data, software, and hardware. In any case, these tools optimize and manage IT assets, help the organization determine what they have, control costs and risks, and improve security. To identify and monitor assets and define the role of users in our proposed structure, some steps were proposed (Figure 2.1).

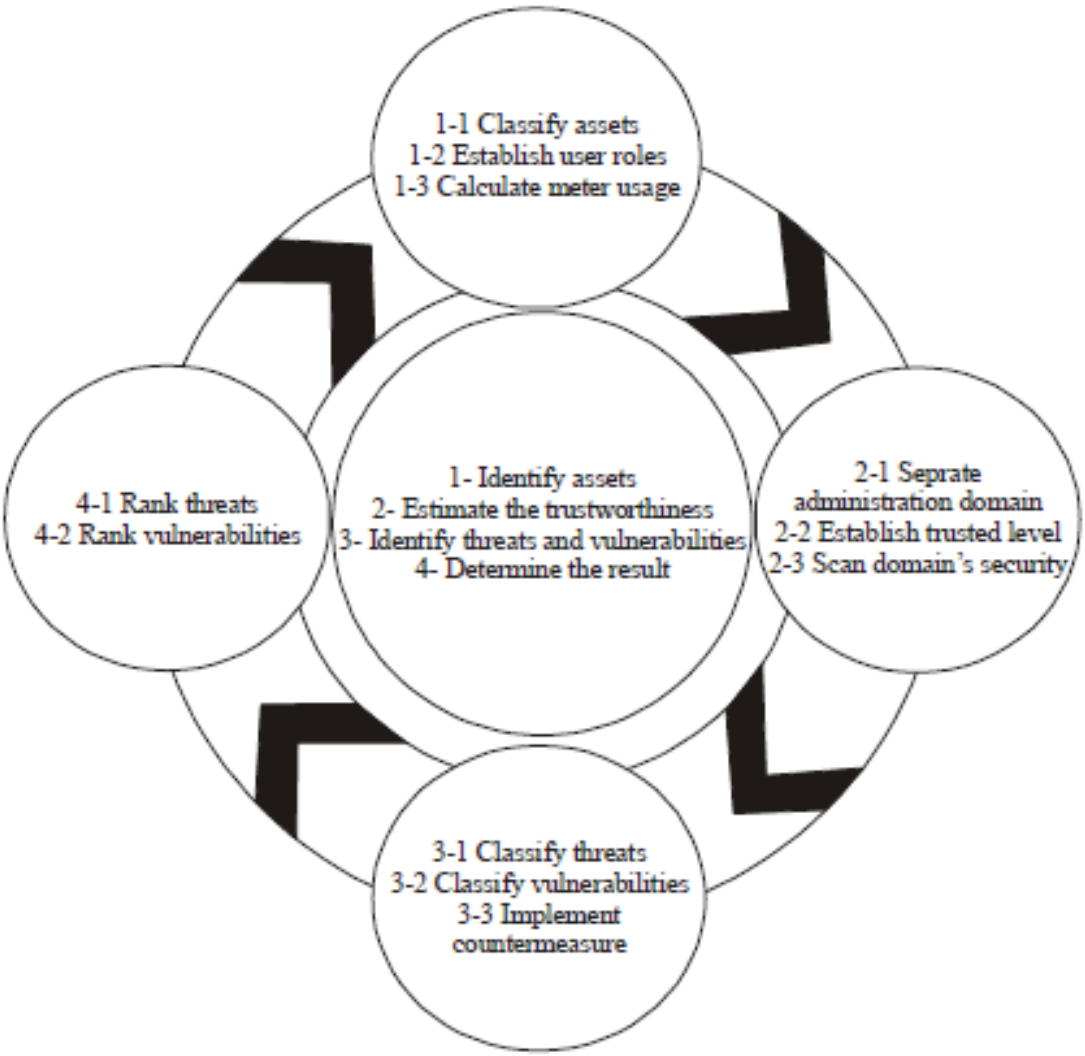


Figure 2.1 - Cloud Computing Threat Modeling Framework by Amini-Jamil

Assess Reliability. Trust is the reliability of distributed systems between two organizations that rely on each other for security. Currently, trust plays an important role in the integration of heterogeneous environments for reliability assessment by ensuring confidentiality, integrity, availability and reliability. To determine and assess trust between different distributed systems, it is necessary to divide administrative areas, identify system vulnerabilities, and establish the level of trust.

Identification of threats and vulnerabilities. As information and communication technologies develop, organizations are increasingly interested in outsourcing their computing resources to virtual domains. In fact, this huge amount of data can be damaged by threats from different resources (actions of employees or malicious attacks by attackers). Therefore, when managing and assessing risks, it is necessary to understand and analyze threats and vulnerabilities as the most important security issues. Thus, effective security classification is necessary to identify and organize threats and 39 vulnerabilities into classes based on the expected impact of attacks, and to develop solutions to prevent effective threats to the system.

Determine the result. Assessing and ranking the severity of threats and vulnerabilities to make an informed decision about what to do to protect the system from malicious influence is the final step of the proposed model. Availability of a threat and vulnerability ranking system is necessary to save time and prevent more serious security threats. The result of this step can be a list or database of threat and vulnerability profiles consisting of sorted security risks.

2.4 Selection of access control subsystem

Access control is typically a policy or procedure that allows, denies, or restricts access to a system. In addition, it can monitor and log all attempts to access the system. Access control can also detect users trying to gain unauthorized access to the system. This is a mechanism that is very important for protection in computer security. There are different access control models, including the most common models: Mandatory access control (MAC), Discretionary access control (DAC), and Role Based Access Control (RBAC). All these models are known as identity-based access control models. In all these access control models, users (subjects) and resources (objects) are identified by unique names. Identification can be done directly or with the help of roles assigned to subjects. These access control techniques are effective in immutable distributed systems where there is a known set of users with a known set of services.

Currently, large open distributed systems are developing very quickly. These include network computing and cloud computing. These systems are similar to virtual organizations with different autonomous domains (branches). Relationships between users and resources are dynamic and more highly specialized in cloud systems. In these systems, users and providers of resources do not belong to the same security domain.

Users are usually identified by their attributes or characteristics, rather than by predefined identification data. In such cases, traditional identity-based access control models are not very effective, and therefore access to the system must be based on attribute-based decisions.

In addition, in the cloud system, autonomous domains have a separate set of security policies. Therefore, the access control mechanism must be flexible to support different types of domains and policies. With the development of large distributed systems, attribute-based access control (ABAC) is becoming increasingly important.

Consider access control methods.

The first way the system ensures the security of its resources and data is by controlling access to resources and the system itself. However, access control is more than just controlling which users (entities) can access computing and network resources. Access control also allows you to manage users, files and other resources. It controls user rights to access files or resources (objects). In access control systems, various steps such as identification, authentication, and authorization are applied before granting access to resources or an object as a whole.

In the early stages of information technologies, researchers and technologists realized the importance of preventing users from interfering with each other's work in shared systems. Various access control models have been developed. The user's personality was the main indicator that allowed users to use the system or its resources. This approach is called Identification Based Access Control (IBAC). However, with the growth of networks and the number of users in them, it was found that IBAC proved to be weak to protect against such a large number of users. Improved access control concepts were introduced to include owner/group/public. IBAC also proved problematic for distributed systems. Managing access to the system and resources has become difficult and vulnerable to errors. A new method known as role-based access control (RBAC) was introduced. Role-based access control defines user access to the system based on their role. The role assigned to a user is primarily based on the principle of least privilege. A role is defined with the fewest permissions or functions necessary to perform a job. Permissions can be added or removed if the privileges for a particular role change. However, problems became apparent when RBAC was extended to administrative domains. And it turned out to be difficult to agree on what privileges to associate with this or that role. Then the attribute-based access control (ABAC) model appeared. In ABAC, access is granted by attributes that the user can provide, such as

date of birth or phone number. However, it is very difficult to agree on the necessary set of characteristics, especially between numerous institutions and organizations. All access control methods are based on user authentication on the site and during requests. Sometimes these methods are called authentication-based access control methods. All these methods require a close connection between the domains. In addition, all of these approaches make it difficult to assign subsets of administrator rights. This leads to the fact that common usage patterns can be implemented by reducing functionality or violating the principle of least privilege.

Damiani E. in his work New paradigms for access control in open environments made an attempt to provide a single framework for the specification and enforcement of ABAC. P. Bonatti in his work A unified framework for regulating access and information release on the web presented a unified framework for formulating and justifying restrictions on access to services and information disclosure based on the relevant attributes of the organization.

Attribute-based access control extends role-based access control with, in general, the following features.

1. Delegation of authority to define attributes.
2. Decentralization of attributes and functions.
3. Interference (intersection) of attributes.

ABAC provides an authorization privacy policy. This allows the organization to maintain its autonomy while cooperating effectively. In addition, it provides automatic trust negotiation that can be checked as needed.

2.5 Selection of the audit subsystem

Cloud computing is very promising for applications in the field of information technology (IT), but for personal users and enterprises, a number of issues related to data storage and deployment of applications in the cloud computing environment have yet to be resolved. Data security is one of the most significant barriers to implementation, followed by issues such as regulatory compliance, privacy, trust and legal issues. Therefore, one of the most important goals is to maintain the security and integrity of data stored in the cloud, given the critical nature of cloud computing and the large volume of complex data being processed. First, users' security concerns should be addressed to make the cloud environment reliable and help users and enterprises adapt to it at scale.

The main challenges in cloud data security include data privacy, data availability, data placement, and secure transmission. Threats, data loss, service interruptions, external malicious attacks and multi-user attack problems are the main security issues in the cloud. Data integrity in the cloud system means maintaining the integrity of the stored information. Data must not be lost or altered by unauthorized users. Cloud providers are trusted to maintain data integrity and accuracy. Data privacy is also an important aspect from the user's point of view as they store their personal or sensitive data in the cloud. An access control strategy is used to ensure data confidentiality. The problem of data privacy can be solved by increasing the reliability of cloud computing. Therefore, the security, integrity and confidentiality of data stored in the cloud must be considered and are important requirements from the user's point of view. To achieve all these goals, it is necessary to develop and implement new methods or techniques.

Data auditing is introduced into cloud computing for secure data storage. An audit is a process of checking the user's data, which can be carried out both by the user himself (the owner of the data) and by a third party auditor. This helps maintain the integrity of data stored in the cloud. The role of the verifier is divided into two parts: the first is a private audit, that is, only the user or owner of the data has the right to check the integrity of the stored data. No one else has the right to query the server for this data. But this leads to an increase in verification operations per user. The second is the possibility of public audit, which allows any person, not only the client, to make a request to the server and perform a check of the authenticity of the data using a Third-Party Auditor (TPA). A TPA is an entity used to act on behalf of a client. She has all the necessary knowledge, capabilities and professional skills that are necessary to perform data integrity verification work. It is important that the TPA effectively checks the cloud data storage without requesting a local copy of the data. He should have zero knowledge of the data stored on the cloud server.

There are three network objects in the cloud environment - the client, the cloud server and the TPA. The client stores data on a server provided by a cloud service provider (CSP). TPA verifies the client's data, periodically checking the integrity of the data upon request and notifies the client of any changes or errors found in the client's data. Figure 2.2 shows the architecture of cloud data storage.

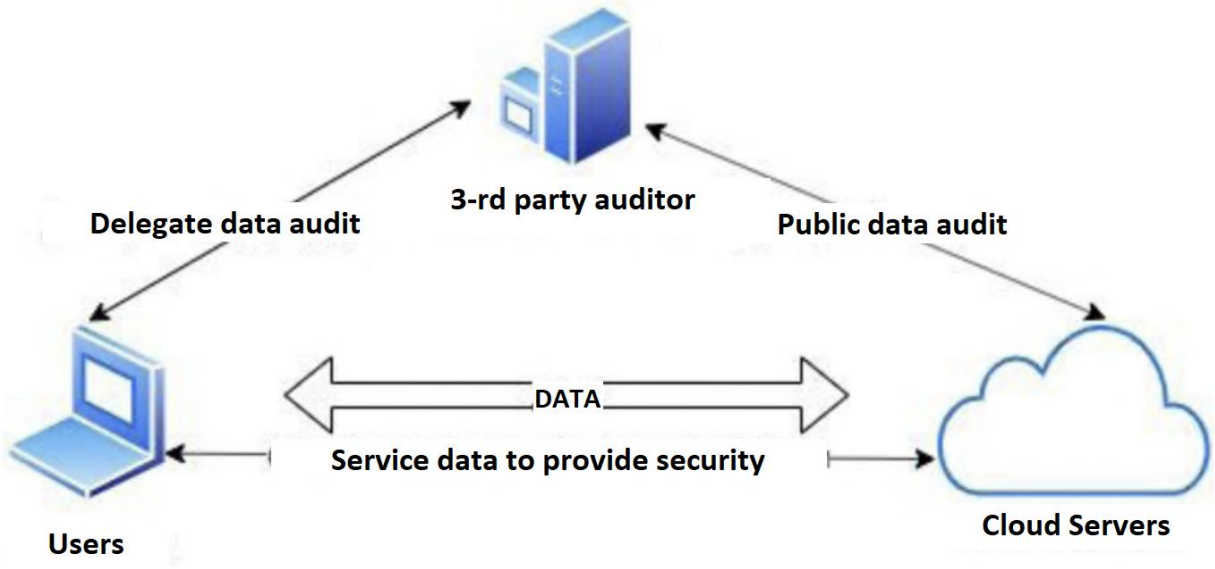


Figure 2.2 - Cloud architecture of data storage

We should take into account various factors such as the method used, public audit support, privacy preservation, data dynamics, and batch auditing. And also whether the integrity and confidentiality of the data stored on the cloud server is maintained or not.

It is necessary to choose an effective public audit protocol that would allow overcoming the limitations imposed by other audit systems.

2.6 Selection of database cryptographic protection subsystem

The storage and processing of sensitive data in a system provided by a third party increases the risk of unauthorized disclosure if the system is compromised by an attacker (who may himself be an agent of that third-party service provider).

One possible solution to this problem is to encrypt the data on the client machine (which is assumed to be trusted) before uploading it to the server, and execute the requests by getting the encrypted data back from the server, decrypting it, and executing the request on the client machine. However, database queries and analytical workloads require much more data to be transferred than necessary, as a large portion of the database is read to execute the query, but the result itself is usually a small aggregated set of data or a convolution of data, such as the sum of item costs.

When analyzing literary sources, three systems were found to solve this question: CryptDB, MONOMI, and Victor Tello's system.

MONOMI builds on previous work on querying encrypted CryptDB databases, and solves some of its problems without introducing new flaws. Therefore, I consider it not appropriate to consider CryptDB, due to the existence of the same ideology, but a better implementation of the system.

MONOMI represents an approach based on separate client / server execution. Using encryption algorithms that allow operations on encrypted data, such as comparisons and groupings. Split request execution allows MONOMI to execute part of a request on the server. For other parts of the request, which cannot be performed on the server at all, MONOMI downloads intermediate results to the client and performs final calculations there. The MONIMI scheme can be seen in Figure 2.3

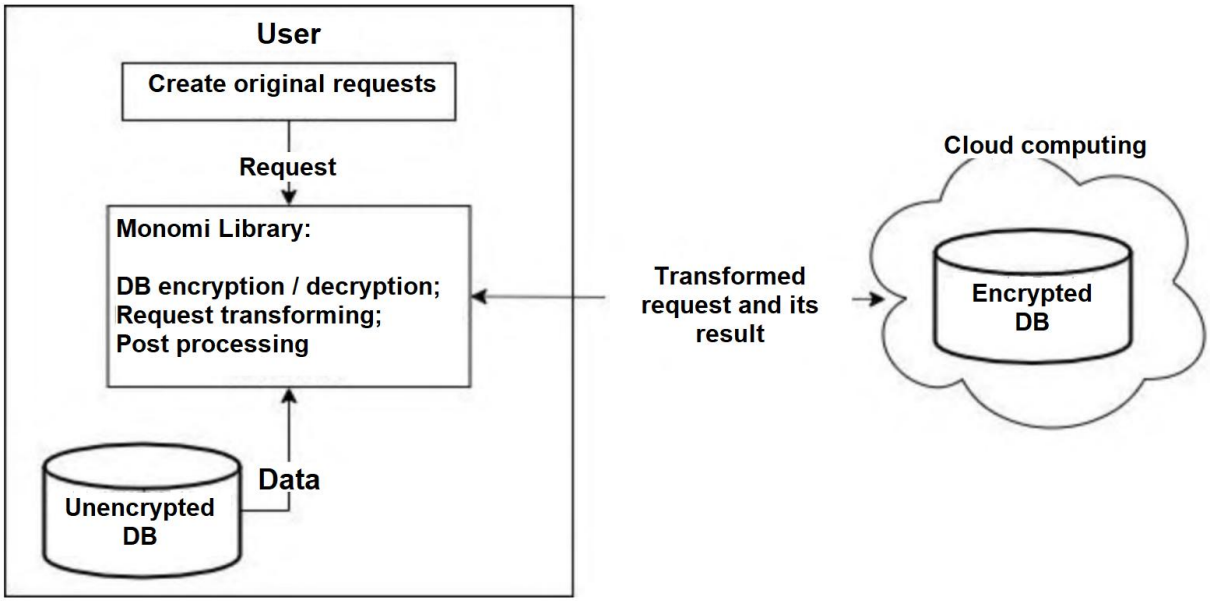


Figure 2.3 - Scheme of MONOMI

MONOMI can be compared to Victor Tello's system, which is schematically depicted in Figure 2.4.

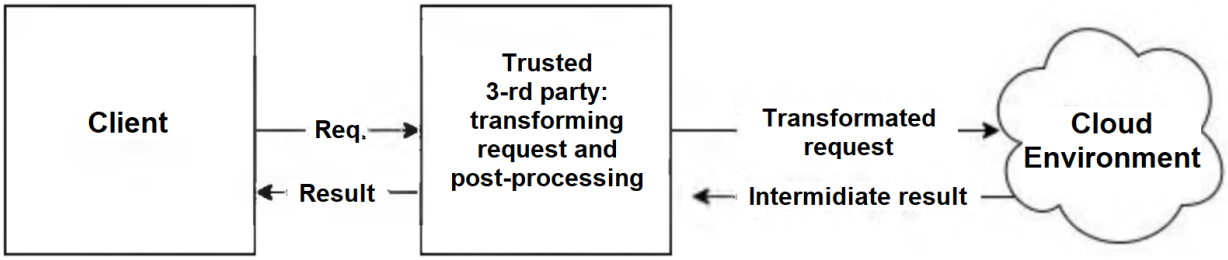


Figure 2.4 - Schematic representation of Victor Tello's system

In MONOMI, a fairly significant part of the calculations is performed on the client side, which negates the significant advantage of using cloud databases, and the

algorithms used in this system to perform operations on encrypted data allow data leakage, namely duplicates, order, and partially clear text.

While Victor Tello's system uses a trusted third party (separate entity) rather than the client to perform all intermediate computations such as: performing intermediate computations, partitioning and indexing the database in encrypted form to the cloud database, and is not allowed data leak.

Therefore, Viktor Tello's system should be used in modern cloud databases.

Conclusions to section 2

In this section, vulnerabilities and threats in the field of cloud services, existing frameworks for threat modeling were considered, and the best of the proposed frameworks was selected according to some criteria.

Threat modeling methodologies can effectively identify and assess security risks in complex systems, allowing system architects to mitigate potential security issues early in the life cycle when they are relatively easy to resolve, and late in the system's life cycle.

The threats and vulnerabilities discussed show that in order to keep information safe, protection against its misuse must be applied.

It was determined which of the existing algorithms, protocols, models it makes sense to use in modern security subsystems of cloud databases.

3 RESEARCH OF SECURITY TECHNIQUES FOR PROTECTING DATA IN CLOUD COMPUTING

In the previous sections, it was made a conclusion about the importance of security techniques for data protection in cloud computing. Based on that analysis we can distinguish the following important security subsystems in cloud computing that will be discussed in this chapter.

- VPN tunnels to protect access to cloud infrastructure.
- Cloud data correctness audit.
- Cryptographic protection of the cloud data storage.

3.1 Using a VPN to secure access to cloud computing

VPN (Virtual Private Network) – is a logical network based on virtual tunnels, created on the top of another communication network. Whereas, even if communications are carried out over public networks using insecure protocols, information exchange will be protected from outsiders using encryption so the data confidentiality and integrity are ensured. Access to such a virtual tunnel should be extremely difficult for all possible active and passive external observers. VPN allows you to combine, for example, several offices of an organization into a single network with secure transmission of information over the Internet.

In cloud computing, a VPN can be used in the following ways - as a Remote Access VPN and as a Site-to-Site VPN.

Remote access VPN is used to provide clients with secure access to cloud resources over the public Internet. This is especially true when you use a public Wi-Fi hotspot or other insecure connection methods to connect to the cloud endpoints on the Internet. Remote access VPN also requires client software to be installed on the client device. This VPN client software interacts with the VPN gateway, which authenticates and authorizes the user, and creates a secure "virtual" tunnel between the local network and the VPN gateway. After successfully completing this procedure, the user gets access to the internal network resources of cloud computing (VM, databases and various cloud services) as if he were connected to the internal network. Example of the remote access VPN connection scheme for AWS is shown in fig. 3.1.

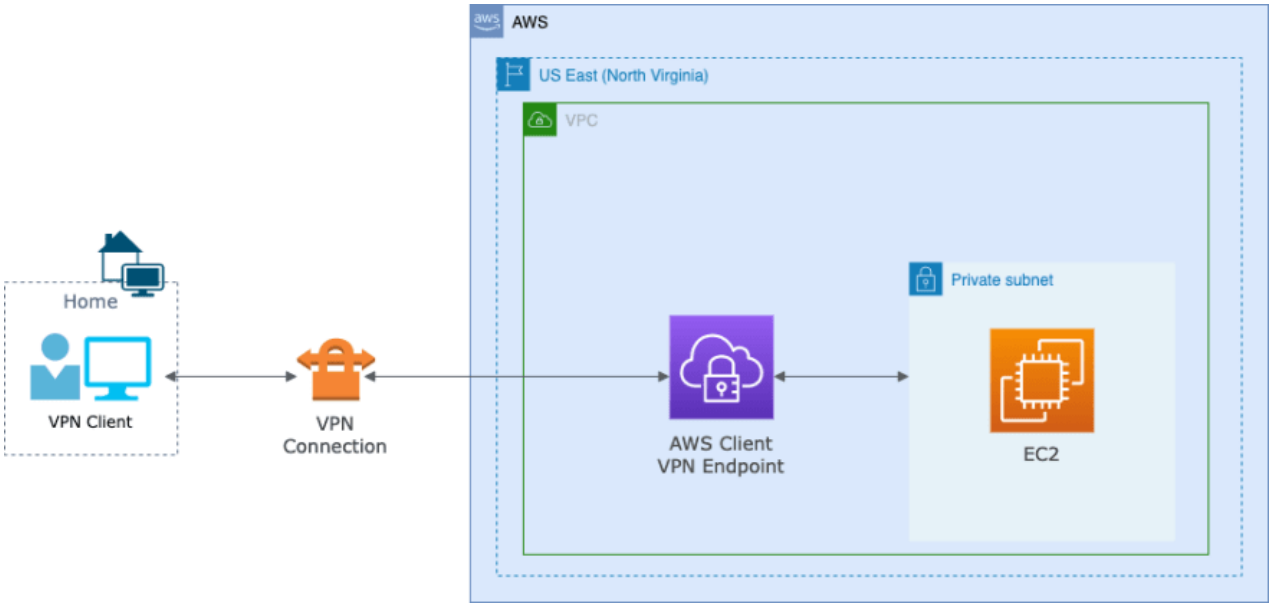


Figure 3.1 - Example of the remote access VPN connection scheme for AWS

A Site-to-Site VPN is used to connect an entire local network in one location to a local network in another location. For example, this type of VPN could be used to make a connection between an on-premises network and Cloud Infrastructure (Figure 3.2).

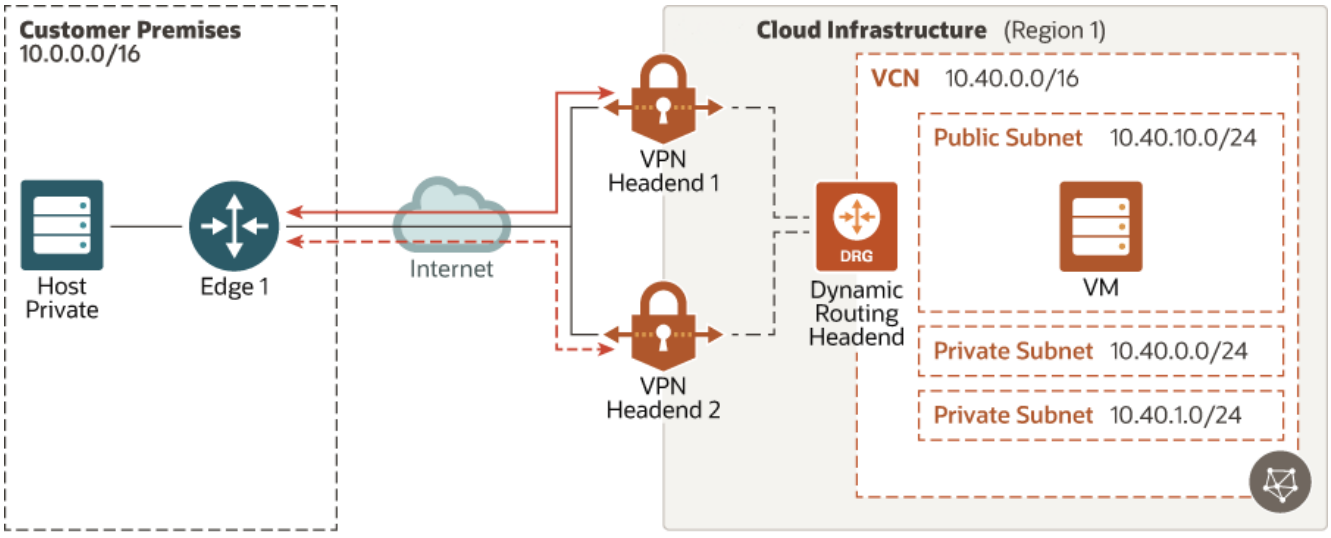


Figure 3.2 - Connecting on-premises network to Cloud Infrastructure using VPN

Such VPN connection between on-premise and cloud infrastructure will have the following advantages:

- your traffic to and from cloud will be encrypted;
- IP address spaces that used in your infrastructure are private and will not expose to outside;
- you can use public internet connection instead of leased lines;
- allowing several users to access cloud resources via a single connection as opposed to multiple connections, a site-to-site VPN lowers the administration burden.

You can use VPN connections also to protect multi-cloud communications. This is a strategy for deploying an IT infrastructure based on multiple cloud providers and platforms, without being tied to a single service provider. As you can see on Figure 3.3 two different VPC from GCP and AWS cloud providers connected through BGP tunnels.

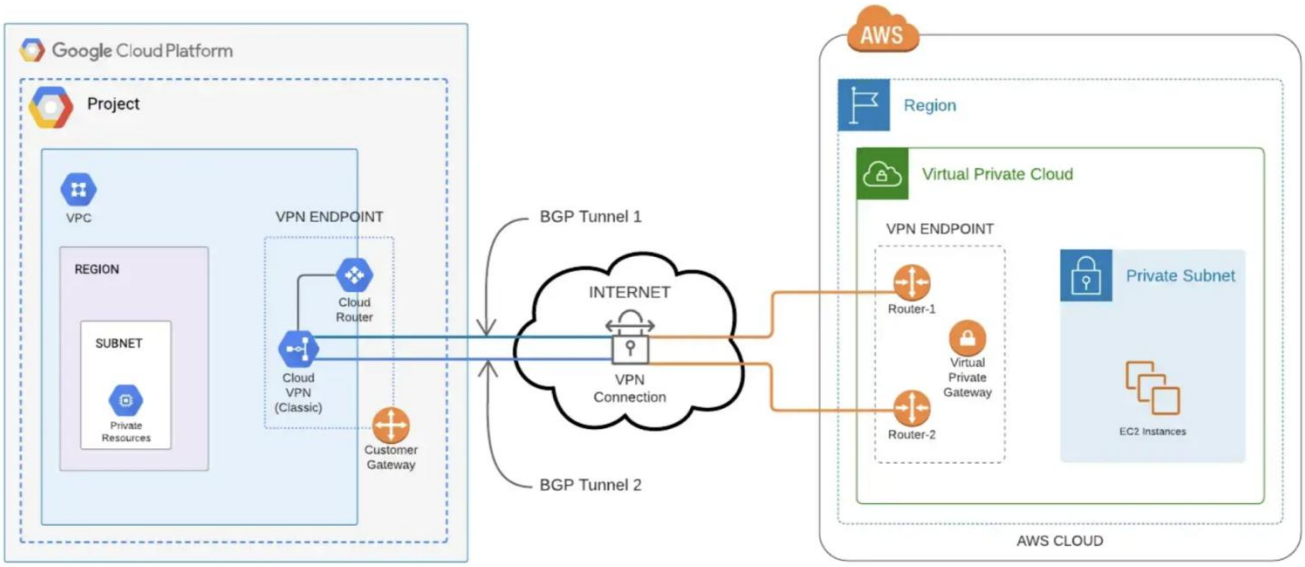


Figure 3.3 - Example of using a VPN to secure a multi-cloud interaction between GCP and AWS

3.2 Cloud data audit subsystem

The chosen Swapnali More system is designed to verify the correctness of cloud data by a third-party auditor. This can be done periodically or by request, without obtaining all data or creating an additional load on users of online cloud environments and on the cloud servers themselves. It ensures the absence of disclosure of data by a third-party auditor during the audit process. It supports the correct storage of data, their integrity and confidentiality.

The proposed scheme consists of three main elements: data owner, cloud server storage and third-party auditor. The owner or user of the data is responsible for splitting the file into blocks, encrypting it using the AES algorithm, generating a SHA-2 hash value for each file, concatenating the hashes, and generating an RSA signature on them. The cloud server is used only for storing encrypted blocks of files. Thus, it has no additional burden of calculating verification evidence. The verification proof here refers to the generation of hashes for encrypted blocks, their concatenation, and the generation of a digital signature for verification. This task is performed by the external auditor himself. When a customer or data owner requests a data audit from a third-party auditor, it immediately requests the encrypted data from the cloud server. After receiving the data, the auditor generates a hash value for each block of encrypted files. It uses the same SHA-2 algorithm as the client. Later, it concatenates these hash values and generates an RSA signature for that file. In the verification process, the signature generated by the auditor and the signature stored by the auditor, which is provided by the user of the data, are compared by the auditor. If they match each other, it means that the data is not corrupted and the data has not been tampered with by outsiders or attackers. If the signatures do not match, this indicates that the integrity of the data has been compromised or tampered with. The results of the data integrity check are provided to the data owner. Figure 3.4 shows the architecture of the proposed audit scheme.

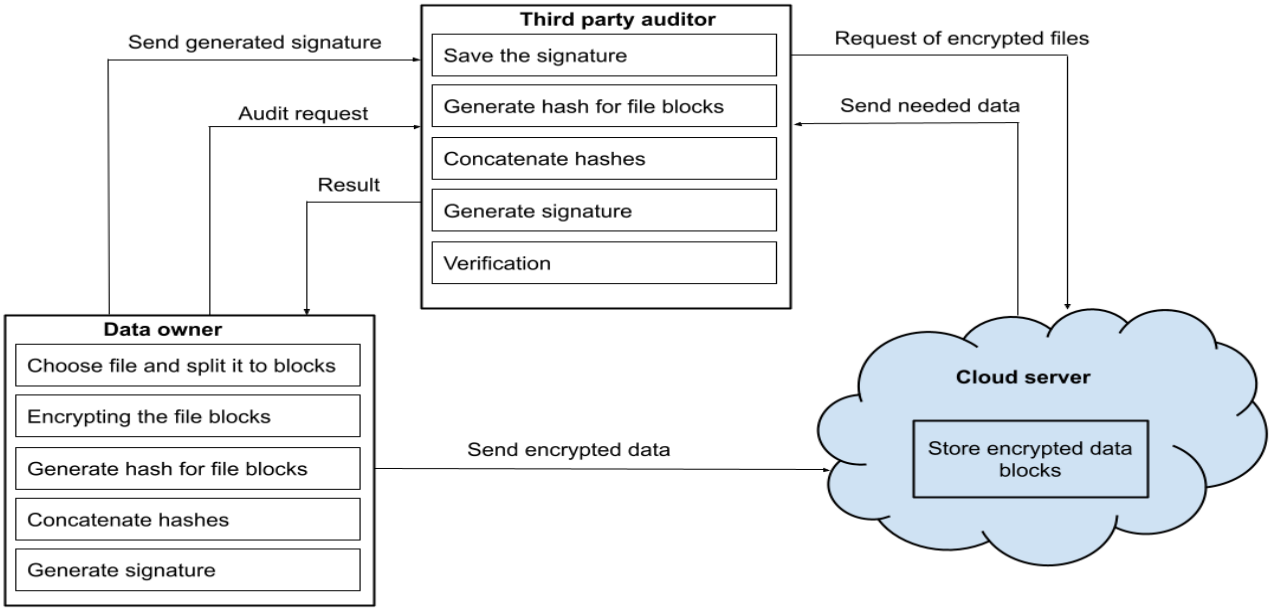


Figure 3.4 - Architecture of the proposed audit scheme

The data owner is an important part of this system. It handles most of the data-related responsibilities. In the proposed audit scheme, the data owner at first logs in and registers on the cloud server and in the third-party auditor service. A new user must first register by filling out the registration form and be active in the system. It will receive a notification of successful registration. If the user is already a system user, it can log in. If the user's login and password exist in the database, they will successfully log in as valid users, otherwise they will receive an error message.

After successful login, the data owner will select the file he or she wants to save on the cloud server. The file selected by him will be divided into several blocks. In order to split the desired file into blocks, the FileSplitter algorithm is used. This algorithm checks whether the file exists or not. If it exists, the file is divided into blocks of a given size, which depends on the size of the file. For example, if the file size is 23 kb, it will be split into 20 kb and 3 kb. Here, in the example, the partition size is set to 20 kb. Next, a strong AES (Advanced Encryption Standard) encryption algorithm is used to ensure data confidentiality. Splitted blocks are now encrypted by the data owner using the AES algorithm. Each block of the file will be encrypted and stored on the client. A copy of the encrypted file will be transferred to a cloud storage. It encrypts 128-bit blocks of data using 128-bit symmetric keys. After encrypting the blocks, the hash value for the blocks is generated separately. For this, the SHA-2 hashing algorithm is used. After the hashes are generated, the hashes for each block are concatenated and an RSA digital signature is placed on it. Digital signatures are used to verify the origin of messages. Later, this signature is sent to a third-party auditor service, where it uses this signature to verify the integrity of the data stored in the cloud storage. The data owner has the right to request a third-party auditor to verify the integrity of the data. Figure 3.5 shows the work of the data owner within this audit scheme.

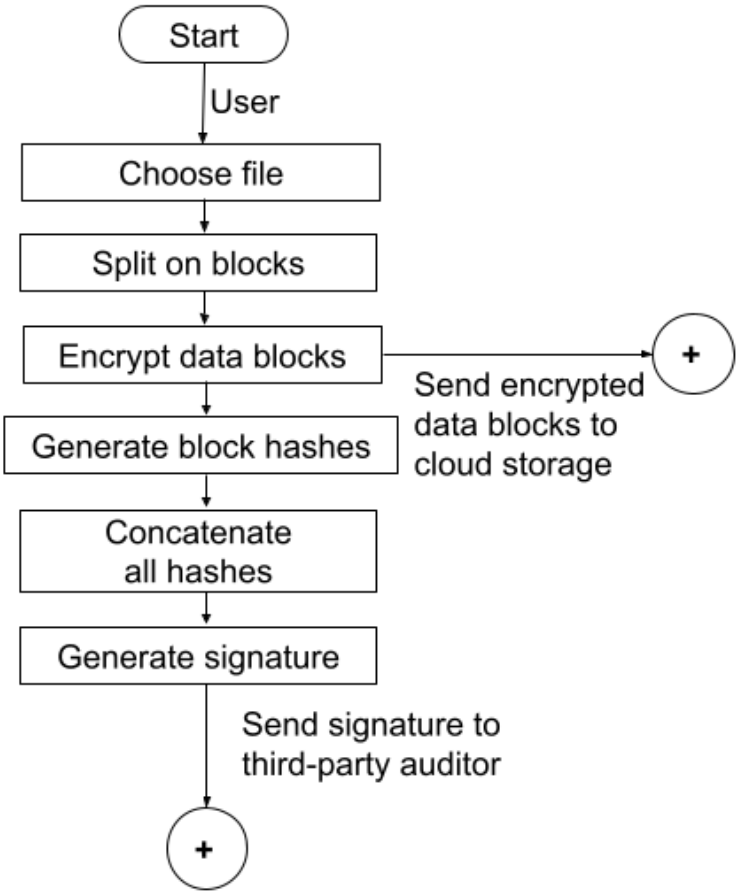


Figure 3.5 - Block diagram of the data owner’s work

The data owner uses cloud storage to store encrypted data. Since the data is stored in encrypted form, the cloud server has no knowledge of the data. As in the case of the transformation of a cloud server into a malicious server or an attack by an external attacker, the data will not be easily retrieved, since it is in encrypted form and the server does not know about the encryption algorithm implemented by the owner of the data.

In this scheme, a third-party auditor is used to perform the data verification task. It conducts data audits either periodically or at the client's request. After receiving a request from a user or data owner to conduct an audit, the third-party auditor begins the audit process. The auditor also stores the signature that was created by the data owner. It follows the same algorithm performed by the data owner, namely, generating a hash for the encrypted data blocks, concatenating them, and generating a signature on them. Later, in the verification process, it compares the two signatures. If they match, it means that data integrity is maintained. Otherwise not supported. This means that the data has not been tampered with or altered. The auditor provides the relevant results to the data

owner. Figure 3.6 shows the work of a third-party auditor within the scope of our audit scheme.

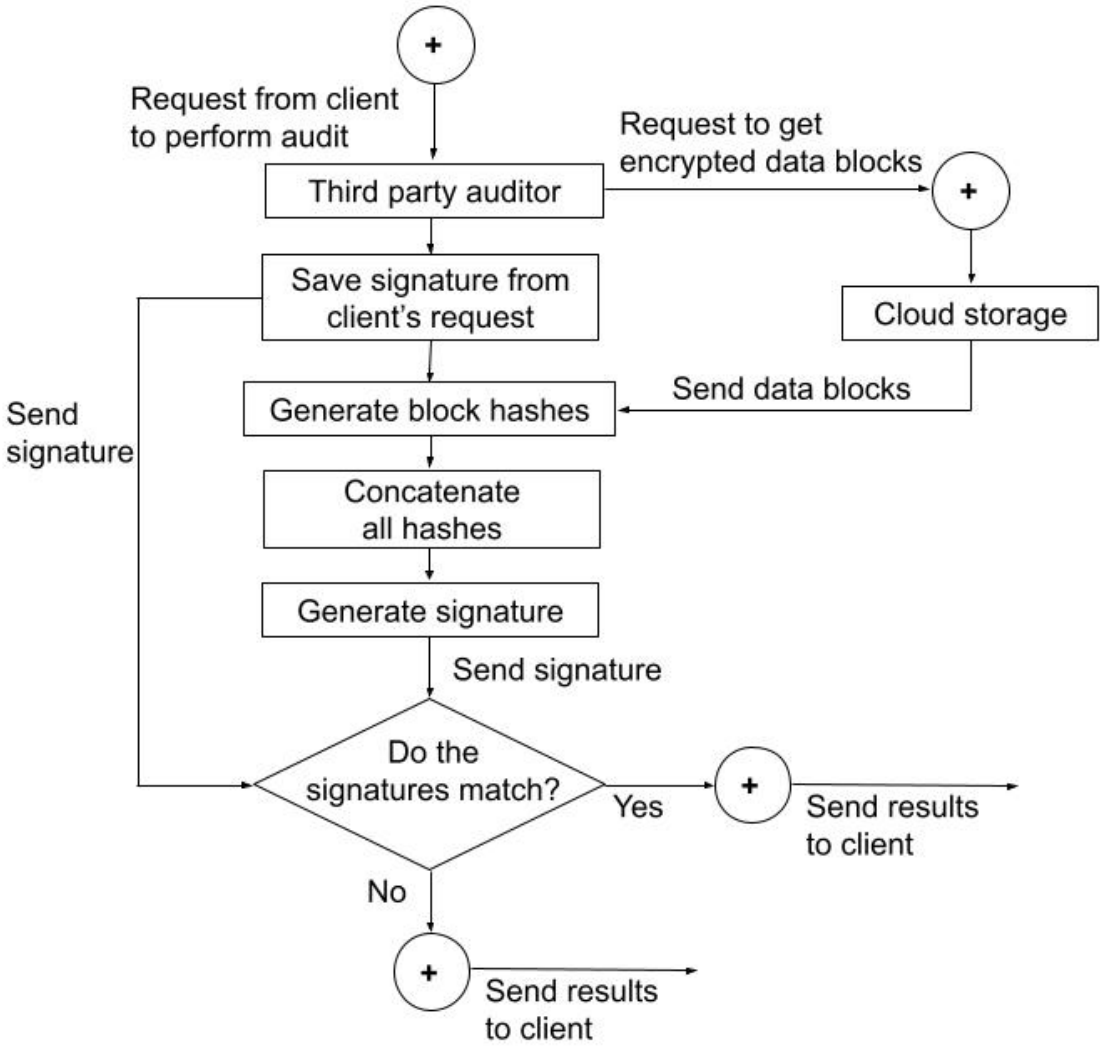


Figure 3.6 - Technological scheme of work of a third-party auditor

3.3 Development of the algorithm for protecting data stored in cloud storage

The disadvantages and problems of data security in cloud systems require the development of new effective algorithms and data protection schemes. The work [12] presents an algorithm that encrypts files uploaded to cloud storages. The integrity and confidentiality of data is ensured not only through encryption, but also through access control with authentication.

Data files uploaded to the cloud are encrypted using the AES algorithm. To increase security, the AES key is encrypted using the RSA algorithm and stored on an internal server outside the cloud.

An authorized user can access data files stored in the cloud and decrypt them.

The model of data protection in cloud storage proposed in [12] has a number of advantages:

- data is sent to the cloud already in encrypted form, in fact this is E2E encryption;
- the decryption key is not saved to the cloud;
- the AES algorithm used is a secure and fast symmetric encryption algorithm. It is fast at both encryption and decryption;
- the ability to frequently change the symmetric key to improve security;
- the AES key used to encrypt the data is RSA-1024 encrypted;
- double authentication is required to decrypt the data. The user must have access rights to the company's server to receive the key and to the cloud storage to download the file.

The cloud data protection model considered in [12] can be improved. In this work, it is proposed to add a third-party audit to the considered data protection scheme in cloud storage. This will improve the level of data protection.

Figure 3.7 shows the data protection model integrated with third-party audit in cloud storage.

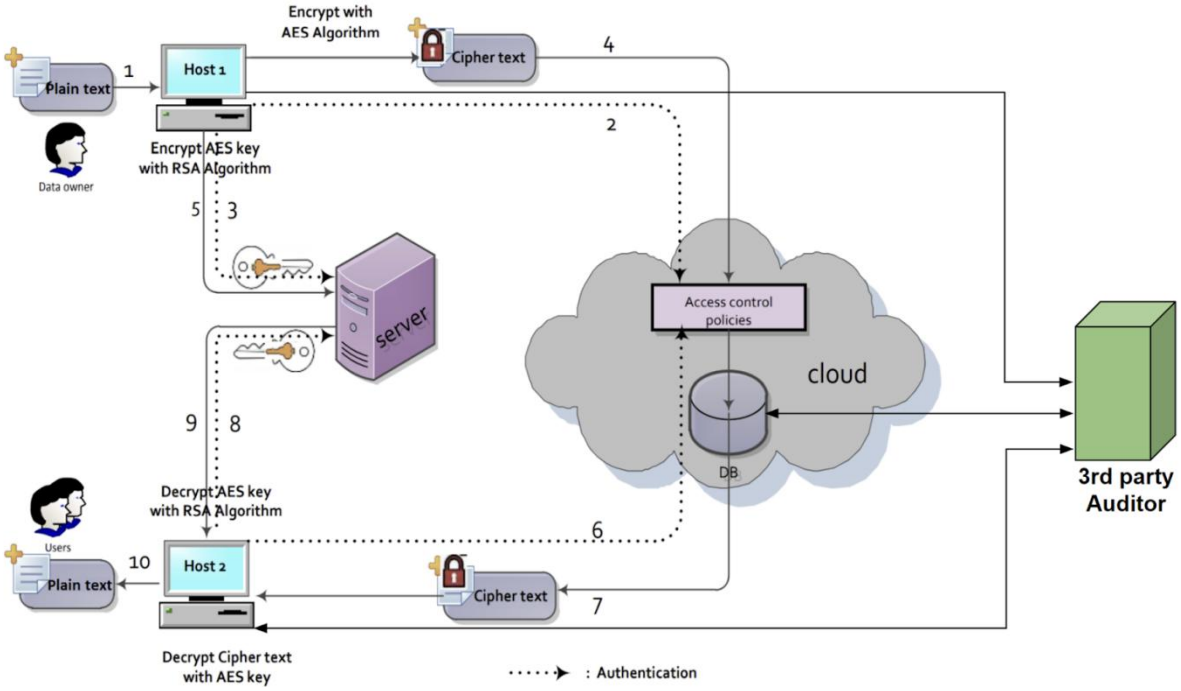


Figure 3.7 - Proposed integrated with third-party audit model to protect data stored in the cloud

The proposed model introduces an additional important element - the 3rd party Auditor, shown in the figure on the right. It also shows the links of its interaction with other elements of the model.

Introducing auditing before sending data to the cloud improves the overall level of data protection. As a result of a third-party audit, by comparing the calculated signature of the data block hashes with their correct signature, you can verify that the data stored in the cloud storage has not been changed.

3.4 Implementation of the data protection algorithm in cloud storage and study of its operation time

When implementing an improved algorithm, you need to find out how the addition of auditing affects the overall performance of the data protection system in the cloud.

Figure 3.8 shows a graph of the performance study results for the original algorithm obtained in [12].

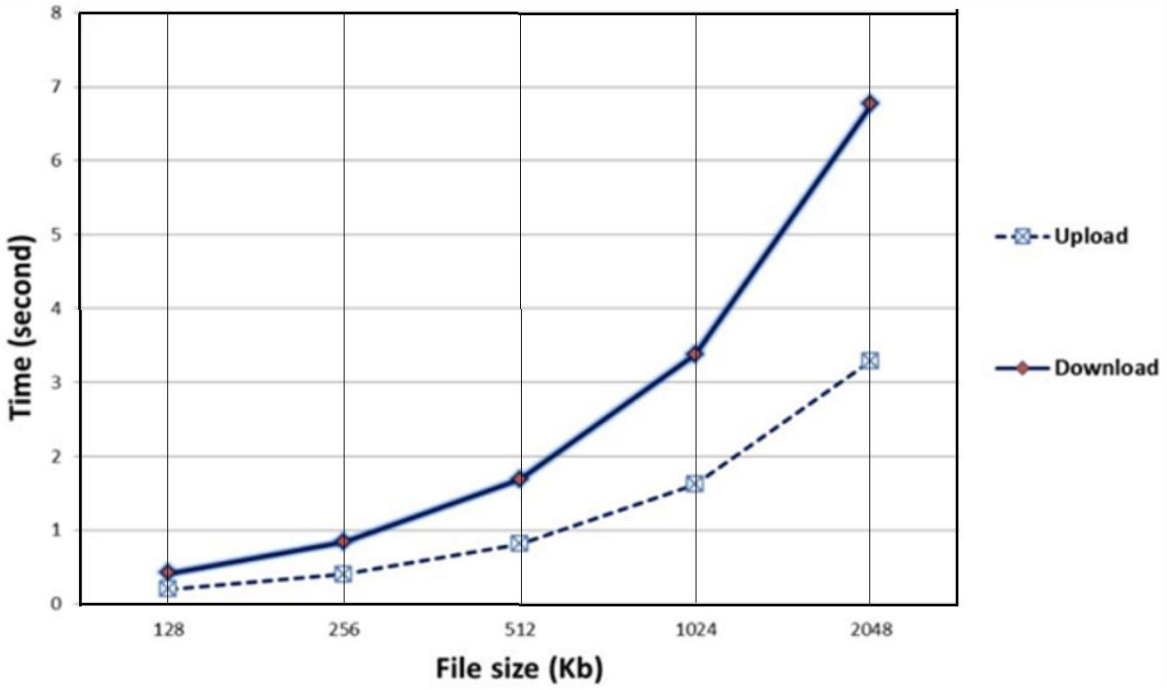


Figure 3.8 - Program execution time for encryption and decryption

Let’s see how our improvements will affect this execution time. In this Master work, a software model for third-party auditing of data uploaded to cloud storage was developed. The programming model of third-party audit was written in Java and is given in the Appendix.

Using a programming model, a study was made of the time of a third-party audit for test data files of various lengths. The partition data block size was 20 kB.

The results of the 3rd party Auditor time cost calculations are shown in Figure 3.9.

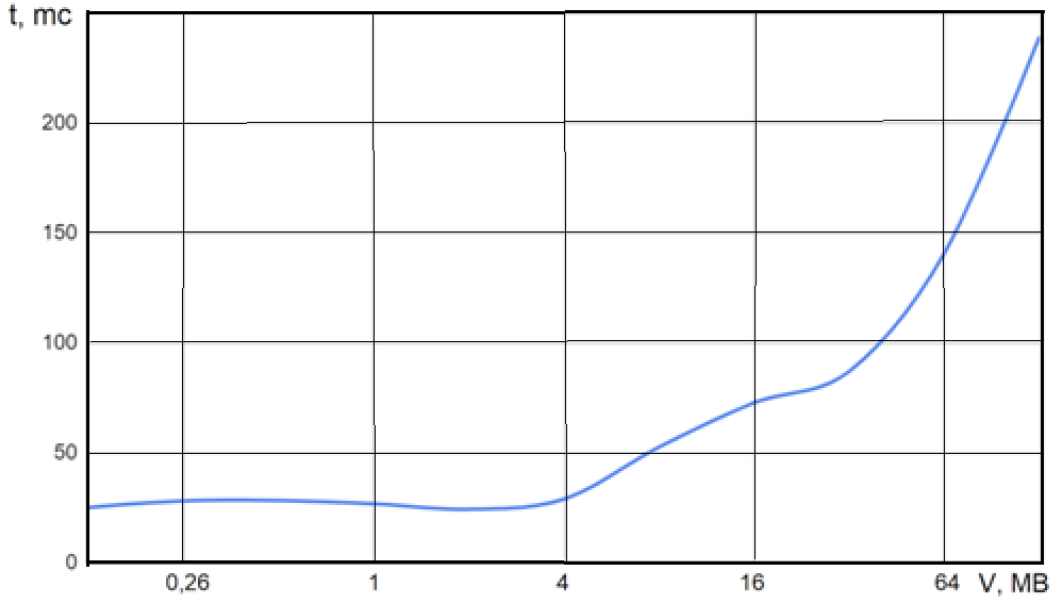


Figure 3.9 - Program execution time for 3rd party Auditor

The analysis of the obtained results shows that adding a third-party audit 3rd party Auditor does not significantly load the work of the original algorithm in terms of time. Thus, the integration of third-party auditing with the underlying algorithm can be used to increase the overall security of data that is uploaded to the cloud storage.

CONCLUSION

In this work, the basic concepts of cloud computing, cloud deployment model, their main properties and architecture were analyzed.

Cloud computing is relevant for use in almost all areas of our lives and various data.

However, cloud service providers do not always put the security of their users' information in the first place, which leads to the appearance of vulnerabilities, which was discussed in the second section. We must build a cloud data security system, taking into account the presence of various vulnerabilities.

The best technological solutions were analyzed and identified for some of the important subsystems that make up the security system of data storage in the cloud, namely:

- for data audit - third party public audit scheme;
- for cryptographic protection of the database - Victor Tello's system;
- VPN tunnels to protect access to cloud infrastructure.

It has also been proposed to combine the cloud storage data protection algorithm with a third-party public audit scheme and store encryption keys separately from the cloud storage. This algorithm was implemented in the form of a Java 11 program. The operation of the algorithm was studied when downloading and uploading files of different lengths.

The obtained results confirm the possibility of practical use of the integrated method. The implementation of the improved method does not require a significant increase in the computing power of the hardware.

The results of the work can be used to build an information protection system applicable to cloud databases by companies of any size, from small startups to large concerns.

REFERENCE

1. Miller R. Who Has the Most Web Servers? .30.12.2022. [Electronic resource]. - Access mode: http://www.datacenterknowledge.com/archives/2009/05/14/whos-got-themost-web-servers/

2. Stephen Watts, Muhammad Raza. SaaS vs PaaS vs IaaS: What's The Difference & How To Choose. 30.12.2022 . [Electronic resource]. - Access mode: https://www.bmc.com/blogs/saas-vs-paas-vs-iaas-whats-the-difference-and-how-to-choose/

3. Bertino E. L. Security for Web Services and Service-Oriented Architectures [Text] / E. L. Bertino // Proceedings of the 2th Annual International Conference on Information Security, New York, USA. - 2012. September. - P. 35-69.

4. Soares L.F.B. Secure user authentication in cloud computing management interfaces [Text] / L.F.B. Soares // Proceedings of the IEEE 32nd International Performance Computing and Communications Conference, San Diego, CA. - 2013. – December. - P. 1-2.

5. Lohman T. DDoS is cloud's security Achilles heel. 30.12.2022. [Electronic resource]. - Access mode: http://www.computerworld.com.au/article/401127/ddos_cloud_security_ac hilles_heel/

6. Massimiliano Rak, Massimo Ficco and Ermanno Battista, Valentina Casola, Nicola Mazzocca. Developing secure cloud applications. [Electronic resource]. - Access mode: https://www.researchgate.net/publication/275640297

7. Hamza Y.A. Cloud computing security: Abuse and nefarious use of cloud computing [Text] / Y.A. Hamza - Int. J. Comput. Eng. Res. - 2013. - 53 p.

8. Fuentes V.T. Enforcing database security on cloud using a trusted third party based model [Text] / V.T. Fuentes // 2438, Theses and Dissertations, ScholarWorks@UARK. - 2017. - 50 p.

9. Cong Wang, Qian Wang and Kui Ren, "Ensuring Data Storage Security in Cloud computing" [Text] / Cong W. // 978-1- 4244 -3876-1/2009 IEEE.

10. Liu Qin, Wang Guojun, et Wu Jie. Secure and privacy preserving keyword searching for cloud storage services. [Text] / Liu Q. // Journal of network and computer applications. – 2012. - Vol. 35, no 3. P. 927-933.

11. Kenneth Hui. Data Encryption in the Cloud. Part 4: AWS, Azure and Google Cloud. [Electronic resource]. - Access mode: https://cloudarchitectmusings.com/2018/03/09/data-encryption-in-the-cloud-part-4-aws-azure-and-google-cloud/

12. Zaid Kariti, Mohamed El Marraki. Applying Encryption Algorithm to Enhance Data Security in Cloud Storage // Engineering Letters. - 2015. – November.-. P.35-38.