UNIVERSITY OF PLOVDIV "PAISII HILENDARSKI"
FACULTY OF MATHEMATICS AND INFORMATICS
DEPARTMENT "COMPUTER TECHNOLOGIES"

ROSEN PETROV HRISTEV

RECOVERY OF INFORMATION ARRAYS IN A CLOUD ENVIRONMENT

ABSTRACT

of dissertation work
for awarding the educational and scientific degree “**Doctor**”

Higher education field: 4. Natural sciences, mathematics and informatics;
Professional field: 4.6. Informatics and computer science;
Doctoral program Informatics

Scientific supervisor: Professor Angel Golev, PhD

Plovdiv, 2023

The dissertation thesis has been discussed and directed for defense at an extended meeting of the Department of Computer Technologies at the Faculty of Mathematics and Informatics of University of Plovdiv Paisii Hilendarski.

The dissertation contains 113 pages. The literature used includes 72 sources - 55 titles in Latin, 3 in Cyrillic and 14 Internet addresses.

The list of author publications consists of 3 titles.

The defense of the dissertation thesis will take place at 22.01.2024 from 11:00 in the Meeting Hall of the new building of the University of Plovdiv Paisii Hilendarski, Plovdiv.

The materials for the defense are available to interested parties at the office of the Faculty of Mathematics and Informatics, new building of Paisii Hilendarski University, room 330, every working day from 8:30 AM to 5:00 PM.

Author: Rosen Petrov Hristev

Title: "Recovery of Information Arrays in a Cloud Environment"

Paisii Hilendarski University Press

Plovdiv, 2023

Content

List of Abbreviations

AFP – Apple Filing Protocol (network file-sharing protocol)

HTTP – Hypertext Transfer Protocol (hypertext transfer protocol)

LDAP – Lightweight Directory Access Protocol (protocol for accessing online directory services)

MICI – Method for Implementing Cloud in existing Infrastructure

NAS – Network-attached storage

NFS – Network File System (network file-sharing protocol)

OCC – ownCloud Console (console interface for ownCloud and NextCloud)

RaaS – Ransomware as a Service (model for distributing ransomware)

RDSC – Recovery Data Stored in Cloud

RVSC – Recovery Versions Stored in Cloud

SaaS – Software as a Service (software delivery model)

SAN – Storage Area Network (data storage network)

SMB – Server Message Block (network resource-sharing protocol)

VPN – Virtual Private Network (virtual private network)

WebDAV – Web Distributed Authoring and Versioning (network protocol operating over HTTP)

DBMS – Database Management System

Thesis Overviw

Impact

The technological development of corporations causes the need of implementing various information technologies to optimize operations. The digitization process, which has been additionally fostered by the Covid-19 pandemic of 2020, makes storage servers one of the commonly used technologies within corporations. Storage servers allow easy data circulation among company's employees through a dedicated IT infrastructure.

This type of IT infrastructute is widely used among a broad size of companies, variating from small and medium sized business to large international corporations. Typically, the data stored in the infrastructure of a company exceeds many times the financial value of the hardware needed to store it.

On the other hand, data is among the vulnerable resources circulating within an IT infrastructure. In addition to the standard threats related data security, such as natural disasters, human errors, hardware and software defects, and others, malicious software developments are continually evolving. In recent years, attacks targeting information arrays have been growing in scale. One of the largest ransomware attacks, the WannaCry attack in May 2017, managed to affect over 200,000 computers in more than 150 countries worldwide. Following the introduction of Ransomware as a Service (RaaS) to the market at the end of 2020, there has been a notable increase in ransomware attacks of over 50% on a global scale.

Standard protection methods, such as data backup, are not always secure enough. Backups can be potentially encrypted in case of poor security management procedures within the corporation. Additionally, there is a risk of information loss between backup and the moment of infection.

Thesis Goals and Tasks

The main objective of this dissertation thesis is to (A) create a method for integrating cloud services into an existing infrastructure, as well as (B) develop methods and tools for recovering deleted information sets among with previous versions of files stored in a cloud infrastructure.

The main objectives are:

1. Research and classification of data threats within IT infrastructures. Subtasks:
   1.1. Investigate standard threats leading to data loss of information arrays;
   1.2. Explore methods of malicious software distribution;
   1.3. Classify malicious software leading to data loss;
2. Study of storage and approaches methods for information arrays within IT infrastructures. Subtasks:
   2.1. Examine standard and cloud-based infrastructures for information array storage and processing within IT infrastructures;
   2.2. Investigate scalability options of cloud infrastructures;
   2.3. Comparative analysis of different methods for storage and processing of information arrays;
3. Analyze the possibilities of recovering deleted data from IT infrastructure. Subtask:

3.1. Analyzing the differences in recovering deleted files stored in standard and cloud infrastructures;
3.2. Approbation of a method to recover deleted data stored in cloud infrastructure;
3.3. Design scripts for automated recovery of deleted data stored in a cloud infrastructure;

4. Analyze the possibilities of recovering previous versions of files. Subtasks:
4.1. Validate a method for recovering previous versions of files stored in cloud infrastructure;
4.2. Design a script for automated recovery of previous versions of files stored in a cloud infrastructure.

Thesis Structure and Size

The dissertation thesis comprises an introduction, three chapters, a conclusion, a list of publications related to the topic, and the validation of results, a bibliography, and has a total length of 113 pages. The thesis also includes an author's contribution statement, a declaration of originality, and prospects for future development.

Chapter 1 discusses multiple storage methods in modern IT infrastructure. It outlines the differences between public and private clouds as well as most common threats to data arrays, leading to data loss.

Chapter 2 explores methods for accessing data arrays within IT infrastructures. It examines key differences in data storage and access between standard and cloud infrastructures. An original MICI method is developed to implement a cloud infrastructure in an existing infrastructure. Furthermore MICI suggests a migration method from standard infrastructure to cloud storage. MICI has been validated in more than 10 real IT infrastructures that vary in size and scope. Some of the organizations where the private cloud has been implemented and the method has been applied are: the Faculty of Mathematics and Informatics at Paisii Hilendarski University, Kaspela University Hospital, Colorado Furniture Factory, WrightPack Printing House Ltd.

Chapter 3 discusses the recovery of deleted and overwritten information arrays stored in cloud infrastructures. In this context, two proprietary methods were developed to recover data circulating in a cloud environment: RDSC and RVSC. RDSC is a method designed to recover deleted data, whitch has been stored in a cloud environment. The work outlines automated scripts to recover data deleted from the private cloud, using RDSC method. RVSC is the second authoring method that was developed and described in Chapter 3. RVSC is designed to recover previous versions of data stored in a cloud environment. Author-created scripts for automating this task are also included in Chapter 3. Both methods methods and scripts are an important tools for data recovery in cloud environments and contribute to the higher security standards security and resilience of data stored in cloud infrastructures.

Summary of dissertation

Chapter I. Data storage and security threats in IT infrastructures

Local area network storage methods

Data storage on a local area network can be realized through a variety of methods and technologies depending on the specific needs and requirements within corporations. The choice of such a storage method is determined by various factors, including the specific needs of the organization and users, budget, security, and data volume. In common cases a combination of different methods seems to be a potential solution to provide storage, security and accessability of data within corporations.

Shared directories typically are considered as a separate repository of shared information. Shared directories are a method of storing data on a local computer network where files and directories can be shared and accessed by different computers on the network. Network protocols are used to access shared directories on the local computer network, allowing computers to communicate and share data. The network protocol depends on the operating system, the needs to which it will be applied, and the network environment. Different protocols have individual characteristics and compatibilities. Some of these protocols are the following: *Server Message Block* (SMB), *Network File System* (NFS), *Apple Filing Protocol* (AFP), and others.

Network Attached Storage (NAS) and Storage Area Network (SAN) are two different data storage methods which are implemented in information technology and computer networks. They have different purposes and characteristics, and are implemented depending on the specific needs and requirements of organizations.

Network Attached Storage (NAS) is a device that is used to store, manage and share files and data on a local computer network. It is typically a server which is hosted on the network and provides file services to users through multiple network protocols. NAS is a specialized device designed to store data and files.

Storage Area Network (SAN) is a type of network architecture designed for centralized data storage and management. It provides high-capacity, high-speed storage where data is stored on separate devices called storage servers. Within SANs data can be shared by multiple servers on the network. SAN is commonly implemented in large organizations which have large volumes of data. The large volumes of data cause a necessity of high performance and storage reliability. SAN is particularly suitable for servers where virtualization and database is implemented.

Cloud Data Storage

Cloud Storage is a method of storing data on remote servers, usually hosted in dedicated data centres and accessible over the internet. This method allows users to store, manage and share data using resources provided by the cloud service provider.

Types of Clouds - Cloud computing is a revolutionary mechanism that has changed the way enterprises purchase and design software. These new methods are fundamentally changing the way data is stored. Nowadays, almost everything is stored in the cloud instead of running programs and storing data on a specific computer. Cloud technologies provide many advantages to their customers such as free services, easy access over the internet, scalability of available resources and more [Vurukonda, 2016].

There are four main types of clouds: public, private, community and hybrid clouds. Public cloud infrastructure is managed by the provider that offers it for public use [Kaur, 2015]. Private cloud infrastructure is only accessible by members of the organization, it can be integrated at third parties only in exchange for granted access [Prachi, 2014]. A hybrid cloud is an infrastructure that is a composite of two or more different cloud infrastructures (private, public) which have unique objects but are interconnected by standard or proprietary technologies allowing in this way the transfer of data or applications between the components [Mladenova, 2011]. A community cloud is somewhat similar to a private cloud, but the infrastructure and computing resources are shared by multiple organizations having common privacy, security, and regulatory considerations [Goyal, 2014].

Public clouds - these are a type of SaaS (Software as a service) services - cloud-based software delivery models that allows end users to access software applications over the Internet. The most common public clouds are Google Drive, Microsoft One Drive, DropBox, Amazon Cloud, Apple iCloud and others. These types of clouds are a common way for storing and sharing non-sensitive data. This said, public clouds do not seem to apply as a recommended option for storing and sharing sensitive and confidential information due to the lack of control over the servers which store the data.

Private clouds - There are various providers offering file storage solutions. The private clouds discussed in this dissertation thesis can be integrated into an IT infrastructure free of charge without an outside organization having access to the files circulating within the infrastructure. Some of the advantages of private clouds are: control, security and encryption, data access, integration and more.

File system structure of private clouds - the private clouds considered in the dissertation store data and files in a file system on the server and DBMS [Odun-Ayo, 2017]. Thus, the server files are stored directly in the server file system, and the meta data along with the data about the files themselves, public shares, and shares between users and others, are stored in a dedicated database. For each user, the following directories may exist:

- files – directory for all resources that are owned by the user. If a directory is shared to a user, the resource owner remains the sharer. If a file or directory is created in a shared resource, then the ownership will be assigned to the first owner of the resource.
- files_trashbin – the directory for resources that which has been deleted by the user. It is important to note that if users delete a resource that is shared with them, the deleted data will not be in the resource owner's profile. Instead the deleted resource will be saved in the directory of the user who deleted it. Deleted files and directories for the private cloud are also automatically deleted from the server if they have been in files_trashbin for more than 30 days. Alternatively, files will be deleted if the user has used more than 50% of their assigned free quota. Resources are deleted using the queue method. The directory can have up to 3 subdirectories: files, keys and versions.
- file_versions – versions for all files owned by the user are saved in the directory. Each version ends with a 12 character suffix that includes the Unix timestamp of

when the version was created. The version is a complete copy of the source file. Resources have no limit on how many versions they can have, and they cannot exceed more than 50% of the user's free space quota. When the user's free space is reduced, the oldest versions are deleted by the queue method.

- files_encryption – the directory is only available if the server's encrypted storage module is enabled. When activated, it will not encrypt the current files which are available in the private cloud, but will encrypt all files that will be uploaded from the time the feature is activated. The directory contains 2 subdirectories: keys and OC_DEFAULT_MODULE.
- uploads – the directory is used for temporary storage of files. When a user uploads a new resource to the cloud, its parts will initially be stored in uploads. Only when the upload of the resource is complete, the resource will be positioned in the correct directory in files. The resource will then be automatically deleted from the uploads directory.

Understanding the file structure of the private cloud would help and facilitate the tasks of recovering deleted and overwritten files, as well as reverting to previous versions of files.

Standard Threats to Data Circulating in an IT Infrastructure

The data stored in an IT infrastructure is one of the most valuable resources, while also being one of the most threatened. The financial value of IT infrastructure typically exceeds the cost of storing and processing data many times over. On the other hand, data is among the vulnerable infrastructure resources, with standard threats to data related to natural disasters, human errors, faulty hardware, software bugs, power supply problems, and others [Bozhikov, 2019].

Malicious Software

Malware is any software that is designed to inflict certain damage on a computer system. Malware types can range from stealing information by modifying it to its complete destruction [Denchev, 2019].

There are different types of attacks through which malware spreads. Among the most common are:

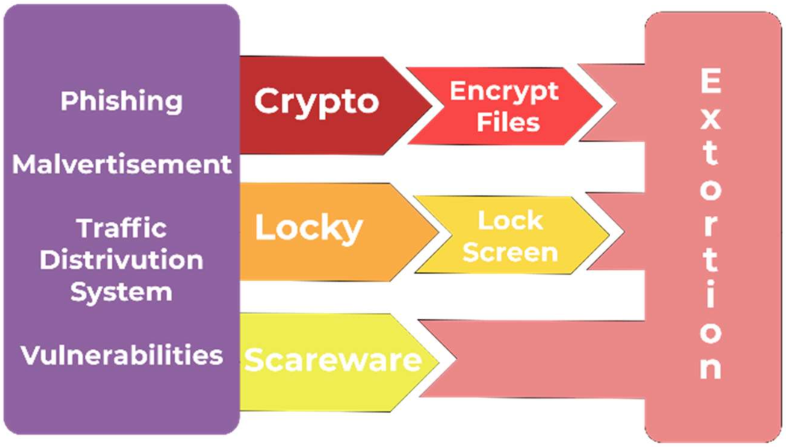
- Phishing attack - this is the most common type of attack and is the sending of electronic messages. It mostly takes advantage of known sources such as email senders in order to deceive the user and gain access to sensitive information [Alkhalil, 2021].
- Using someone else's identity - the attacking entity impersonates another person. This type of attacks are more successful when they are intended for a specific victim and the victim is researched. Thus, it is easier to deceive by sending seemingly credible-looking emails from a user with whom the victim communicated regularly [Altwairqi, 2019].
- Brute Force - this type of attack is applied to poorly secured networks. There are numerous bots which constantly scan public addresses for the presence of running services. Most commonly, standard ports described in IANA are scanned and

multiple login attempts are made with password and user dictionaries when an open port is present [Knudsen, 2011].

- Social engineering - attacks of this type rely on user manipulation. The malicious actions start with gathering information about the victim from corporate websites and publications, or even communicating with the victim. Once enough information has been gathered and a strategy for the attack has been made, the attack itself is resorted to. It often takes advantage of identified weaknesses in the system or manipulating a user who has access to the infrastructure [Salahdine, 2019].
- Environment Intrusion – Attacks of this type are usually characterized by searching for various types of vulnerabilities in the system in order to gain unauthorized access [Stiawan, 2017].

Computer viruses cover a wide range of damage which depends mainly on the type of virus attacking. Cryptoviruses as a family of viruses are undoubtedly the most damaging to files nowadays. Infection with them is usually by executing files containing malicious code. Their main purpose is to encrypt data or devices and demand a ransom in order to regain access to the data [Tailor, 2017].

The peak of the Covid-19 pandemic was a prerequisite for many companies to downgrade the security of their IT infrastructures to allow employees to quickly go into remote mode of work this lead to an increase in ransomware attacks [Venkatesha, 2021]. One of the main reasons for this is the availability of a new type of service, Ransomware as a Service (RaaS), which is due in late 2020 [Alwashali, 2021]. As a consequence, there has been a rise in cryptovirus attacks by over 50% globally.



Фигура 1 Видове криптовируси

Crypto-ransomware - the crypto-ransomware type is considered very dangerous because it encrypts the victim's files, making it impossible to access them without a valid decryption key [Aurangzeb, 2017].

Locker-ransomware - locks the victim's system and displays a login page. The victim must pay a ransom to obtain a password to unlock the system [Richardson, 2017]. Locker-ransomware is considered less dangerous as the attack can often be resolved by rebooting the system into safe mode.

Scareware - this ransomware poses no real danger to its victim. Its main function is to scare its victim, who should pay a ransom [Hrıstev, 2022a]. Scareware attack is often launched through pop-ups that appear on the user's screen, warning him that his computer or files are infected and then offering a "solution to the problem" [Kiru, 2019].

Results and Conclusions

Storage servers are one of the most common in IT infrastructures. Thanks to them, the data circulating in the infrastructure can be easily shared among the company's employees.

From the review and analysis of data storage systems, the following conclusions can be synthesized:

- Standard data storage methods such as shared directories, NAS and SAN servers are not flexible enough.
- Public clouds are a proper solution for storing non-confidential information.
- The available storage space in public clouds can be easily scaled.
- Private clouds are a proper storage solution for confidential and corporate data.
- The disadvantage of private clouds can be their difficulty to scale and the need for technically trained personnel to maintain them.
- Some of the lorgest hacking attacks leading to data loss have been implemented using cryptoviruses.

The following results were achieved in Chapter 1:

1. Methods of storing information sets in an IT infrastructure are researched.
2. Explored and compared some of the most common providers of public cloud services of the SaaS type and private clouds.
3. The scalability of private clouds and the file structure in which they store and process data are explored.
4. Common threats leading to data loss in IT infrastructures are analyzed.
5. Standard malware distribution methods are studied and three types of ransomware are classified.

Chapter II. Access to data stored in IT infrastructures

Chapter 2 of this dissertation thesis discusses various methods of accessing the data that circulates in a modern IT infrastructure. A comparison is made between standard data storage methods and data storage in cloud infrastructures. An author's method is discussed and described, by which it is possible to integrate cloud storage of data sets into an existing infrastructure.

Access to data stored on standard infrastructures

Standard data storage methods are limited to distributing the data across a single local computer network. This leads to a number of limitations that have hampered many organizations with the announcement of the Covid-19 pandemic in early 2020. The move to remote working mode has accelerated the digitization of work many times over, which, in turn, has provoked numerous corporations globally to very quickly reorganize their established work mode. This forced IT specialists to lower the criteria for IT security.

Meanwhile, malware developers became increasingly inventive to achieve their goals. For example, many malware attacks, including cryptoviruses, have begun to use additional scans for system vulnerabilities in addition to standard phishing attacks. A clear example of this is that in 2022 alone, the share of companies that have implemented VPNs to access private networks has increased by 2%, reaching a total of 95% for organizations worldwide. CVE's zero day exploits reports show that nearly 700 VPN-related vulnerabilities have been reported since 1999, of which about 38% have been reported since the start of the Covid-19 pandemic [CVE, 2023]. The situation is similar for one of the most common protocols when it comes to sharing network resources, with 256 vulnerabilities since 1999, of which 58 have been announced since the beginning of 2020 to date.

With shared directories and NAS servers, data is stored at the file level and is most often accessed through protocols such as Samba and NFS [Gibson, 1997].

One of the main problems is that data storage through shared directories, NAS and SANs rely heavily on backups and archives of data when data sets need to be restored [Preston, 2022]. More than 70% of small businesses that are major users of shared directories and NAS do not have a written plan, strategy, and procedure for recovering from data loss. Even with backup retention plans written out, periodic checks are almost never made to see if the consistency of the data in the backups is present and how recoverable the data is from the backups.

Data stored in shared directories and NAS devices is most often accessed as mounted disks over the network [Gobioff, 1997].

Access data stored in public and private clouds

It is rare to find public clouds which does not support file access through a web-based application. This type of file access is not always convenient or appropriate as it would create additional steps for the user when using a file stored in cloud storage [Grance, 2011]. Apart from accessing the data through a web browser, some of the public cloud providers also provide synchronization software. Due to these, the data stored in the cloud is synchronized with a directory on the user's local machine. This downloads a copy of the file to the user's workstation. When the user makes a change to a file or creates a new one, then the syncing client will upload the changes, and with most providers the old data will be saved as a different, previous version of the file.

The considered private clouds offer three main options by which the data stored on them can be accessed: through a web browser, through a synchronization client, and through the use of WebDav.

WebDAV (Web Distributed Authoring and Versioning) as an extension was announced in 1996 by Jim Laithead and is actually an upgrade of the HTTP protocol and is based on it. WebDAV additionally allows users to manipulate data sets directly on the server [Whitehead, 1999]. The protocol is implemented by the private clouds considered and through it disks can be mounted in the operating system for users to see and work with as with file servers [Rieger, 2011].

In Linux and macOS based operating systems the protocol works as standard and can be used without limitations. In some Windows based operating systems malfunctions are encountered, including additional changes that need to be made in some older versions. For example, in Windows 7 and Windows 8, editing is required, including in the operating system registries. Windows-based operating systems can communicate over the WebDav protocol through a standard server called WebClient, allowing the creation and editing of Internet-based files [Whitehead, 1998].

Differences in Storing and Accessing Data in Standard Infrastructures and in Public and Private Clouds

Data circulating in standard infrastructures is primarily accessed through protocols such as Samba and NFS. Some NAS equipment manufacturers in recent years have been focusing not only on these two main protocols but also on synchronizing clients and protocols that allow virtualization clusters and other similar software to have access to NAS devices.

Public clouds, on the other hand, are accessed primarily through a web browser or sync client. This may cause difficulties for users under certain conditions. For example, add-ons implemented for online editing of electronic office documents do not support full functionalities. Additionally, information overwrite situation may occur in case a file is downloaded to be processed with an offline version of an office suite. Data in this case can be recovered by comparing previous versions of the file. This requires manual intervention by the user, which can also cause inconsistency of the data stored in the files.

Private clouds support both data-handling functions – a web-based interface and a synchronizing client. However, unlike most SaaS providers in public clouds, they also offer the option to work with files through WebDav. Accessing the cloud via WebDav may sometimes cause difficulties, which are thoroughly discussed in subsection 2.2 of the dissertation. With the proper selection of private cloud infrastructure, it is possible to enhance it with a hybrid solution, providing users with the convenience of working with standard methods of storing information and the security of having data stored in a private cloud.

Table 1 Differences in Data Access in Standard Infrastructures and Cloud Environment

Technology	Infrastructure	Primary Data Access Methods	Advantages	Disadvantages
Shared Directories	File	Samba, NFS	Easy file handling	• Significant reduction in security if data is not circulated in a private network • Data is not protected
NAS	File	Samba, NFS, Synchronization Client	Easy file handling	• Significant reduction in security if data is not circulated in a private network • Data is not protected
SAN	Block	Samba, NFS, FiberChannel, iSCSI	Easy file handling	• Significant reduction in security if data is not circulated in a private network

				• Data is not protected
Public Cloud	Cloud	Web browser, Synchronization Client	Complex structure	• Data is protected with double deletion and file versions • Data can be accessed from public networks without compromising security
Private Cloud	Cloud	Web browser, Synchronization Client, WebDav	Complex structure	• Data is protected with double deletion and file versions • Data can be accessed from public networks without compromising security

Author's Method for Implementing Cloud into an Existing IT Infrastructure (MICI)

In order to implement a cloud into an existing IT infrastructure using the MICI method (Method for Implementing Cloud in existing Infrastructure), the first step is to choose the type of cloud to be integrated into the infrastructure. This can be done based on the specifics of the data expected to be stored in the cloud infrastructure.

Cloud services offer the possibility of file recovery through integrated functionalities such as double-deletion of files and previous file versions. To enable the recovery of information arrays, the available space in the infrastructure must be calculated in advance. The required server space can be calculated using the following formula:

$$DS = D*(1+Cr),$$

where: DS - disk space; D - data; Cr – coefficient increased data size, as a result of the encryption algorithm depending on the type of cryptovirus, where Cr > 1.0.

The examined private clouds can easily be upgraded to comply with ISO 27001 standard - Information Security Management Systems. In this scenario, the server's information arrays will be stored in encrypted form, which would additionally increase the size of the data in unencrypted form by up to 35% [Nextcloud, Encryption Configuration, 2023]. To calculate the required disk space for storing data in encrypted form, the formula above would assume the following format:

$$DS= D*(Ea+Ea*Cr),$$

where: DS - disk space; D - data; Ea - coefficient for increased file size, as a result of activation of a module for encrypted data storage, where Ea > 1.0; Cr - coefficient increased data size, as a result of the encryption algorithm depending on the type of cryptovirus, where Cr > 1.0.

The structures by which cloud systems store their data differ from the file structures in standard data storage methods. They have a subdirectory structure that is not directly visible to users, and it is also tied to the DBMS used by the cloud. The directory tree may vary depending on the cloud, but the server's file system has a directory for each user with access to the cloud infrastructure. The directory is usually named after the user or has a unique identifier in LDAP authentication. Within the user's tree, there are directories with names such as:

- files – location where all files owned by the user are stored;
- files_trashbin – all files deleted by the user;
- file_versions – a directory where all different versions of files created by a specific user are stored.

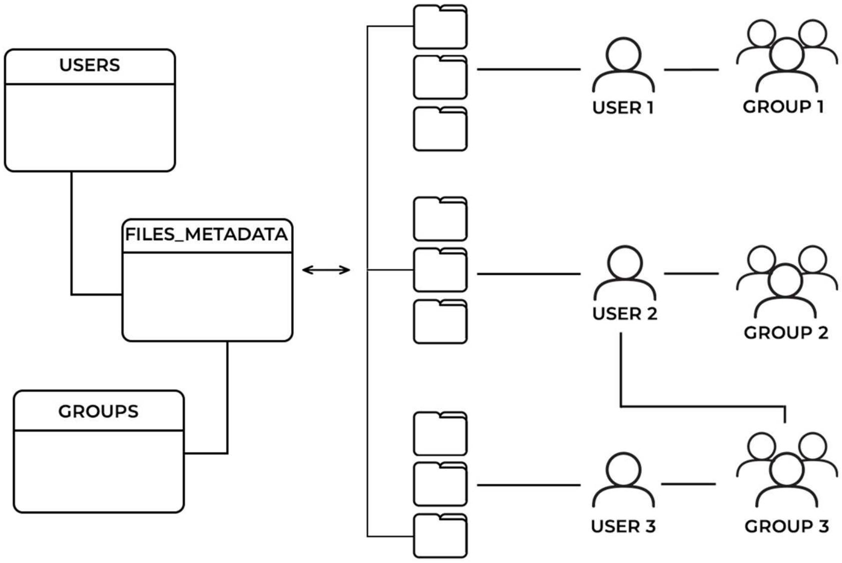


Figure 2 File Structure and Database in Cloud Data Storage

A quota can usually be set for the user's root directory. There is a difference between quotas in public and private clouds. In public clouds, it can be assumed that a user's quota is actually the files they have created or stored in the private cloud, and it doesn't matter whether this is in shared space from another user or is in their own directory. When a user deletes a resource on public clouds, if it was shared with him, the sharing will be removed and the file will be discoverable on the disk of the user who owns the resource. The space for the resource has already been used by the user's quota, and this operation will not further reduce the user's available free space [Hristev, 2021].

From another perspective, in private clouds, if a user uploads a resource to a shared directory, users do not consume their own quota but rather that of the user who owns the shared directory. In practice, ownership of the resource is assigned to the user who owns the shared directory. Metadata (such as temporary files, caches, and encryption keys) is taken from the quotas of both users, with a default setting that specifies they cannot exceed 10% of their respective quotas.

With quotas in private clouds, it is advisable to create primary users for the system where all files are stored and shared with other users. The purpose of this is that when working with the private cloud, the ownership of the files should always be on the primary users. This will allow that when an employee leaves the company, their user can be easily deleted from the system without deleting their files (except those in "Deleted Files"). This eliminates the need to transfer ownership of files, which can be done from the console or the cloud web interface. In addition, the primary users storing the files also contribute to proper disk space management, as it is possible to configure quotas on the users used by employees in order to limit the use of the private cloud not for its intended purpose.

For public clouds, the amount of space a user is recommended to have can be calculated using the formula:

$$US = 1{,}5 * UF,$$

where: US - user space; UF - resources owned by the user.

For private clouds where no encrypted file storage module has been added, the formula for calculating a user's quota would be as follows:

$$US = 2.5 * FA,$$

where: US - user space (quota); FA - resources to which the user has access.

For private clouds that are configured to store data encrypted on the server, the formula would be:

$$US = Ea * 2.5 * FA,$$

where: US - user space (quota); Ea - coefficient for increased file size, as a result of activation of a module for encrypted data storage, where Ea > 1.0; FA - resources to which the user has access.

In summary, it can be concluded that the key steps for implementing a private cloud in an existing IT infrastructure using the MICI (Method for Implementing Cloud in existing Infrastructure) method are as follows:

1. Selection of cloud type;
2. Calculation of the required parameters for the cloud infrastructure - the disks can be calculated using the formulas: DS = D*(1+Cr) or DS= D*(Ea+Ea*Cr);
3. Configuration of the selected cloud;
4. Transfer existing data to the cloud;
5. Adding users and creating a hierarchy - user quotas can be calculated using the formulas: US = 1.5*UF, US = 2.5*FA or US= Ea*2.5*FA;
6. Sharing information with new users;
7. Configuring workstation access to the cloud.

The proprietary MICI method is thoroughly described in subsection 2.4 of the dissertation.

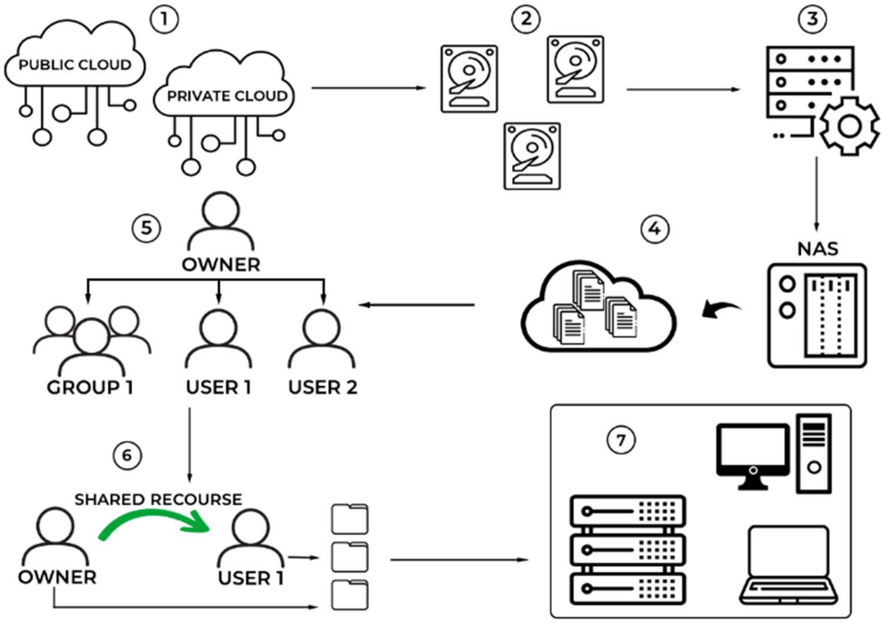


Figure 3 Method for Implementing Cloud in Existing Infrastructure

The author's MICI method has been applied in more than 10 real IT infrastructures that vary in size and areas of operation [Hristev, 2021]. The largest organization that has integrated a private cloud using the method and switched from NAS to NextCloud has over 300 workstations. The method has also been implemented to migrate small companies with less than 10 workstations using shared directories to a private cloud. Some of the organizations where the data has been successfully migrated to a private cloud are: the Faculty of Mathematics and Informatics at University of Plovdiv Paisii Hilendarski, Kaspela University Hospital, Colorado Furniture Factory, WrightPack Printing House Ltd. and others.

The literature review did not find any other similar methods describing the steps for moving from a standard data storage infrastructure to cloud storage of information stored in IT infrastructures.

Results and Conclusions

Transitioning to remote work as a result of the Covid-19 pandemic prompted many corporations worldwide to quickly reorganize their established work routines. At the same time, malware developers became increasingly inventive to achieve their goals.

From **Chapter 2** of the dissertation, the following conclusions can be synthesized:

- In standard information processing infrastructures, data is mainly distributed in a local computer network.
- With the announcement of the Covid-19 pandemic in early 2020, reports of zero-day exploits on the main services facilitating remote access to corporate private networks significantly increased.
- While standard information storage and processing methods use a file system to store data, cloud infrastructures use a combination of file systems and a database.
- Older operating systems and application programs may not support the new network protocols used by clouds to access the data stored on them.

- Cloud infrastructures support different and flexible ways to access data.
- Moving from a standard data storage infrastructure to a cloud infrastructure can be a complex and complicated task.

In **Chapter 2**, the following results have been achieved:

1. Data access methods in cloud infrastructures are analyzed.
2. Approaches that allow older operating systems such as Microsoft Windows 7 and Microsoft Windows 8 to work with data stored in cloud infrastructures are defined and described.
3. A hybrid model of working with cloud infrastructures is analyzed to reduce the network load on the infrastructure.
4. An original MICI method is created to migrate from standard infrastructure to cloud data storage. The developed method has integrated clouds in more than 10 IT infrastructures, varying in size and scope.
5. The methods for calculation of minimum required available resources in cloud infrastructure allowing recovery of information sets after loss are defined.

Chapter III. Recovery of Deleted and Overwritten Data in IT Infrastructure

Chapter 3 of the dissertation thesis discusses options for recovering data that has been deleted or a new version of a file has been saved.

Data Recovery in Standard Infrastructures

In infrastructures that use standard data storage methods such as shared directories, NAS, and SANs, data is typically stored in different types of file systems determined by the specifics of the infrastructure. When operating on hard disk drives, file systems store data in a fragmented form. Since it is not clear in advance what the size of a file will be, and it is quite possible that it will increase in size, the data is written to the first free sector on the physical disk. This means that a file is split into pieces that are written to different locations on the disk platter. Data fragmentation causes significant slowdown in operations such as reading and modification of existing files. For this reason, when the system is not under load, the vast majority of file systems automatically defragment the data or actually arrange the data sequentially, which would speed up the file reading process.

When a resource is deleted, until it goes through a defragmentation process, the freed sectors on the hard disk are marked as free. In file system terms, this means that they are sectors on which new information can be written. In standard computer operation, files are constantly being generated, such as various caches and others, that would be written immediately to the free sectors.

In standard data storage methods, attempts to recover information can be made using deep scanning software for hard disks. These programs conduct a sector-by-sector check of the disk, attempting to recover information for a given resource if metadata for. The file is found. It is sufficient for information to be overwritten on one of the sectors of the hard disk where information about a particular resource is available to compromise the integrity of that resource. Usually, this can happen even during the computer operates.

Even if the infrastructure has implemented backups using the 3-2-1 method, which is currently one of the most secure, data from the last backup until the system infection will still be lost. Another significant drawback of this data storage method, which can be noted, is that the majority of companies do not have scripted action scenarios for data loss, and where they exist, resources are rarely allocated to check the integrity of the archive and the ability to restore the system to a functional state from the archive.

Recovery of Deleted Data in Cloud Infrastructures

Clouds have built-in functionalities such as double deletion of files that can be used to recover deleted information. Considering the nature of cloud operations, the disk space occupied by deleted files is no more than 50% of the quota for a given user. The default period for private clouds is 30 days. The rule is that the first entered in the deleted files will be the first deleted. The required disk space and user quotas can be calculated using the formulas from subsection 2.4 of the dissertation thesis.

RDSC - Author's method for recovering deleted data stored in a cloud environment - it can be concluded that the author's RDSC method for recovering information arrays after data deletion is as follows [Hrıstev, 2022a]:

1. Identifying the compromised link in the infrastructure

   • In the case of a Ransomware attack, the first step is to determine how many workstations in the infrastructure are compromised and infected with crypt ransomware. Depending on the attack, this is usually one or a group of computers from a specific department. It is not excluded that there might be infected computers in multiple such groups.

2. Isolating the compromised links

   • In the case of a rogue employee, the employee's access to the system must be restricted.

   • In the event of a malware attack, once the infected computers are identified, they must be isolated. Omitting this step would cause the method to cycle because after recovering the files through the private cloud, the compromised devices would re-encrypt the recovered files.

3. Cleaning the compromised devices

   • In the case of malware, the compromised devices need to be cleaned from the crypto virus. Companies developing antivirus software invest more resources in updates to clean computer systems from such threats. Once the virus is identified, references can be made online to find products that can detect and clean the infection. If no tool is found to remove the problem, there is always the option to reinstall the infected machines.

4. Recovering deleted files from "Deleted Files"

   • If the deleted data was personal to the user, they will retain their structure and be restored to the locations from which they were deleted.

   • If the deleted files were shared for the user, they will be restored as personal to the user, preserving their original structure.

5. Checking the cloud structure: the structure needs to be checked, and if the files were deleted due to a virus intrusion, it should be verified if there are encrypted files uploaded by the virus at the original location of the files. If such files exist, they should be deleted once from the storage accessible to users and a second time from the "Deleted Files".
6. Restoring files to their original locations and configuring access if necessary.

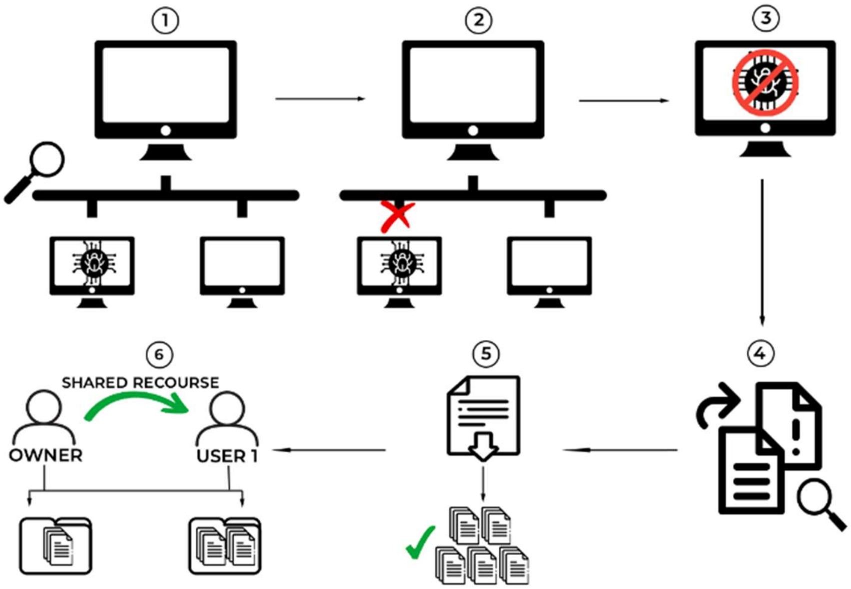


Figure 4 Author's Method for Recovering Deleted Files Stored in a Cloud Environment

Thanks to the author's method for recovering deleted data sets RDSC (Recovery Data Stored in Cloud) were recovered data after infection with cryptoviruses in controlled environments, namely:

1. User data from a controlled Windows workstation after infection with CERBER [Hristev, 2022b];
2. User data from a controlled Linux-based Debian workstation after infection with GonnaCry [Hrıstev, 2022a];
3. Recovery of a WordPress site and MySQL database running on a controlled Linux Debian server after infection with GonnaCry [Golev, 2022];
4. Recovery of a .Net Core application and archives of databases on Microsoft SQL Server running on a controlled Windows Server 2019 after infection with Sodinokibi [Hristev, 2023];
5. Multiple recoveries of various data deleted by unscrupulous employees in different companies.

The literature review did not reveal any other described methods for recovering deleted data stored in cloud infrastructure.

Scripts for automated recovery of deleted data stored in a private cloud using the RDSC method - it is possible to observe a slowdown in the performance of the web-based part of the application [Hristev, 2023]. This, in turn, would lead to difficulties in recovering the encrypted files and deleting the encrypted files from the encrypted files and would take a long time to recover. The NextCloud and OwnCloud private clouds have a complex structure, with all resources available to the user being further described in the server database in addition to the server file structure [Hristev, 2023]. For ease of management, both clouds have an OCC [Nextcloud, OCC command, 2023] in addition to a web-based interface, through which some of the functions are accessible from the shell. This can greatly ease their administration.

Despite numerous requests from users to develop functionality for OCC to recover files from Deleted Files, none has yet been developed [Hristev, 2023]. A script has been created that enables automated file recovery using the RDSC method. The script described in subsection 3.2.6 of the thesis, after slight modifications that are dependent on the type of cloud, will allow recovery of previous versions of files stored in public clouds. In order to implement the restore, administrative access to the servers on which the public cloud is installed and running is required.

Enhanced scripts for automated recovery of deleted data stored in a private cloud using the RDSC method - The script described in subsection 3.2.6 will restore all files located in the "Deleted Files" of the user. In some cases, this could lead to the recovery of unnecessary information, which users would subsequently need to manually delete. For this reason, the script has been revised, and a new improved version for automated data recovery has been proposed. In this version, the improved script additionally takes date ranges as an argument. The script would search for files to recover only within the pre-defined date range.

Recovery of overwritten data

The private clouds and some of the public cloud service providers considered have built-in file version control functionality that can be used when a file has compromised integrity or has been overwritten by rogue users of the system. In addition, there are also cryptoviruses that do not delete files on the system, but actually encrypt only part of the file metadata without deleting the source. In such situations, the information sets can be recovered through version control.

RVSC – proprietary method for recovering overwritten data stored in a cloud environment - clouds retain various versions for the files stored on them. By default, there are no set parameters regarding how far back versions of the respective files should be retained. The limitation on versions is based on quota limits, with the size of stored versions not exceeding 50% of the user's quota's free space [Nextcloud, File Versions, 2023].

The proprietary RVSC (Recovery Versions Stored in Cloud) method for recovering overwritten data stored in a cloud environment is as follows:

1. Identifying the compromised link in the infrastructure.
    - In the case of a user error, it is necessary to determine whether the data overwrite was intentional or accidental. If the overwrite was accidental, proceed to step 4.
    - In the case of data overwrite due to a ransomware attack, it is first necessary to identify how many workstations in the infrastructure are compromised and infected with crypt ransomware. Depending on the attack, this is usually one or a group of computers from a specific department. It is not excluded that infected computers may exist in more than one department.
2. Isolating the compromised links.

- In the case of a malicious employee, access to the system must be restricted initially.
- In the event of a malware attack, once the infected computers are identified, they should be isolated. Skipping this step would lead to a cyclic process in the method because after restoring files through the private cloud, compromised devices would encrypt the recovered versions again.

3. Cleaning the compromised devices – in the case of malware, the compromised devices need to be cleaned from the virus.
4. Identifying the overwritten files.
5. Downloading previous versions of the overwritten files and verifying the data integrity in the downloaded versions
6. Restoring the overwritten files to their original locations and restoring access to them if the files were shared.

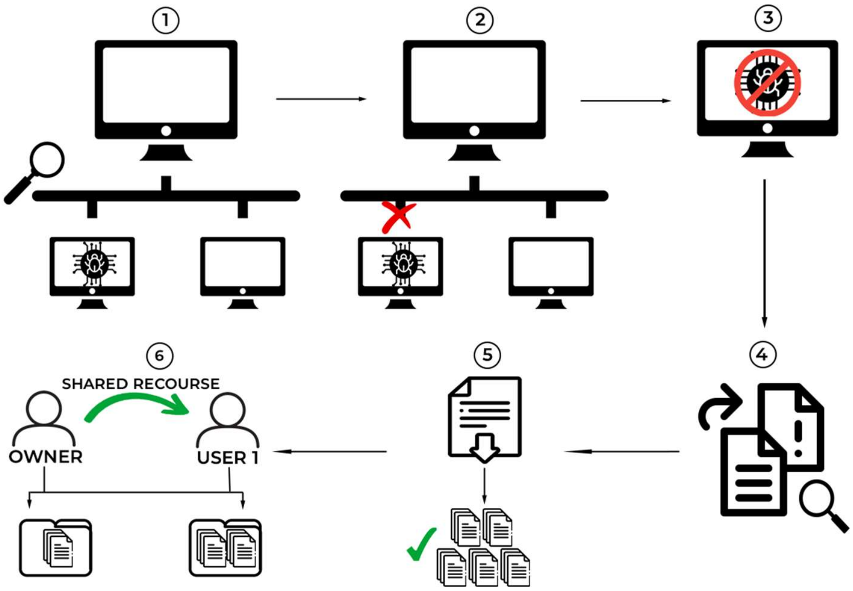


Figure 5 Author's Method for Recovering Overwritten Versions of Files Synchronized with Cloud Infrastructure

Thanks to the RVSC proprietary method, user data that has been overwritten due to negligence or lack of knowledge by employees in various organizations has been successfully recovered multiple times. The recovery method can also be applied to restore data in cases of power outages and loss of information in specific files.

The literature review did not identify any other described methods for recovering overwritten data stored in cloud infrastructure.

Script for automated recovery of previous file versions stored in a private cloud using the RVSC method - recovering overwritten data can be a labor-intensive task since version recovery occurs file by file, and there are no implemented functionalities for such actions on multiple files simultaneously. Additionally, it is recommended that files not be directly restored; rather, they should first be downloaded and their data verified. A script has been developed to enable automated file recovery using the RVSC method. The script takes time and user arguments, restoring the closest version of all files for the respective user. To complete the recovery using the RDVC method, the files need to be moved to their original locations, and if necessary, access to them should be configured. The script, described in subpoint 3.3.2 of the dissertation, will allow the recovery of previous file versions stored in public clouds after slight modifications depending on the cloud type. Administrative access to the servers where the public cloud is installed and operational is required to implement the recovery.

Results and Conclusions

The financial value of data stored in IT infrastructure by far exceeds the value of the hardware required for their storage. Data, in turn, represents one of the vulnerable assets of organizations. In addition to the standard threats for data integrity, there is a continuous evolution and development of malware by malicious software developers. Over the past few years, attacks targeting data arrays have been increasing both, in frequency and scale. Since the introduction of Ransomware as a Service (RaaS) to the market at the end of 2020, there has been an observed increase of over 50% in ransomware attacks globally.

From **Chapter 3** of the dissertation, the following conclusions can be synthesized:

- Deep-scanning software for hard disk drives does not always succeed in recovering lost information due to the specific nature of file system operations.
- Creating backups of information is not always flexible enough, as they can also be deleted or compromised.
- In the majority of corporations, there are no prescribed action scenarios for data loss, and where they exist, resources are rarely allocated to check the integrity of the archive and the system's recovery capabilities from the archive to a functional state.
- In standard data storage infrastructures, recovering files with altered information is nearly impossible.
- Cloud infrastructures have built-in functionalities that can assist in recovering deleted or modified information in data arrays.
- Private clouds can be easily configured to comply with ISO 27001 standards.

In **Chapter 3**, the following results have been achieved:

1. Approaches for recovering data arrays stored in standard and cloud infrastructures have been researched and analyzed.
2. An author’s method, RDSC, has been developed for recovering deleted data stored in a cloud environment. Using the RDSC method, the following data recoveries were successfully executed:
   - User data stored on a Windows-based workstation after being infected with the CERBER ransomware;
   - User data stored on a Linux Debian-based workstation after being infected with the GonnaCry ransomware;
   - Application server and database server on a Linux Debian server after being infected with the GonnaCry ransomware;

   - Application server and database server on Windows Server 2019 after being infected with the Sodinokibi ransomware;
   - Multiple recoveries of various data deleted by unscrupulous employees in companies of various types and sizes.

3. Author’s scripts for automated recovery of deleted data arrays using the RDSC method have been created. The scripts have been improved by adding a time interval for recovery.
4. An author’s method, RVSC, has been developed for recovering previous versions of files stored in a cloud environment. Using the RVSC method, user data that has been overwritten due to negligence or lack of knowledge by employees in various types and sizes of organizations has been repeatedly recovered.
5. An author’s script for recovering previous versions of files using the RVSC method with a choice of a time range has been created.

Summary

As a result of the research conducted in the dissertation thesis, the following results have been achieved:

- A survey of methods and techniques for the storage and processing of data arrays in contemporary IT infrastructures has been conducted.
- Various types of threats to data stored in IT infrastructures have been examined.
- The possibilities for recovering deleted data and previous versions of files have been analyzed.

An author method, MICI, has been developed for moving from standard infrastructure to cloud data storage. The method has been integrated into more than 10 IT infrastructures. Some of the organizations where data has been successfully transferred to private clouds include the Faculty of Mathematics and Informatics at University of Plovdiv Paisii Hilendarski, University Hospital Kaspela, Colorado Furniture Factory, RightPack Printing House Ltd., and others. A hybrid model of working with cloud infrastructures has been analyzed, aiming to reduce network load on the infrastructure.

Common threats leading to data loss in IT infrastructures have been analyzed. Standard methods of malware propagation have been investigated, and three types of exploiting malware have been classified.

An author method, RDSC, has been created, and based on this method, a script has been developed to automate the recovery of deleted and overwritten data in private clouds. It has been improved by adding time criteria for the data. Using the RDSC method, data has been successfully recovered:

- User data from a controlled Windows workstation after infection with CERBER;
- User data from a controlled Linux-based Debian workstation after infection with GonnaCry;
- Recovery of a WordPress site and MySQL database running on a controlled Linux Debian server after infection with GonnaCry;

- Recovery of a .Net Core application and archives of databases on Microsoft SQL Server running on a controlled Windows Server 2019 after infection with Sodinokibi,
- Multiple recoveries of various data deleted by unscrupulous employees in different companies.

An author method, RVSC, has been developed. Based on this method, custom scripts have been implemented to automate the recovery of previous versions of data stored in private clouds by specifying time criteria. Using this method, user data that has been overwritten due to negligence or lack of knowledge by employees in organizations of various types and sizes, where a private cloud is integrated, has been successfully recovered multiple times.

Future work

With the emergence of private clouds and the growing sizes of data circulating in IT infrastructures, the extensively developed custom scripts for automated data recovery will face increasingly wide application. A logical continuation of the foundations laid in the dissertation thesis would be:

1. Expanding the automated recovery scripts based on the RDSC and RVSC methods with metadata verification in the databases of private clouds. This would enable the recovery of data to their initial locations while preserving access rights and resource sharing.
2. Creating modules for NextCloud and ownCloud that allow the recovery of multiple sets of data using the RDSC and RVSC methods through the web-based interface of private clouds.

Primary Contributions of the Thesis

The main contributions in the dissertation are as follows:

I. An analysis of the primary types of threats leading to the loss of information arrays in modern IT infrastructures has been conducted.

II. The development of the MICI method, an author’s approach facilitating the transition from a standard infrastructure to cloud data storage.

III. An author's RDSC method has been developed for the recovery of deleted and overwritten data stored in a cloud infrastructure. Scripts have been created to automate the method.

IV. An author's RVSC method has been developed for restoring previous versions of files stored in cloud infrastructure. Scripts have been created to automate the method.

The relationships between contributions, goals, tasks, the place of description in the dissertation thesis and the publications made are as follows:

Contribution	Goal	Task	Chapter	Publications
I	B	1.	1	3
II	A	2.	2	1
III	B	3	3	2, 3
IV	B	4	3	

Publications

1. R. Hristev, M. Veselinova, ICT for Cyber Security in Business, *IOP Conf. Ser.: Mater. Sci. Eng. 1099 012035*, 2021, ISSN (Online): 1757-899X, doi:10.1088/1757-899X/1099/1/012035,https://iopscience.iop.org/article/10.1088/1757-899X/1099/1/012035/meta
2. A. Golev, R. Hristev, M. Veselinova, K. Kolev, Crypto-ransomware Attacks on Linux Servers: A Data Recovery Method, *International Journal of Differential Equations and Applications*, Vol. 21, No. 2, 2022, pages: 19-29, ISSN (Print): 1311-2872; ISSN (Online): 1314-6084, https://www.ijpam.eu/en/index.php/ijdea/article/view/6002/283
3. R. Hrıstev, M. Veselınova, K. Kolev, Ransomware Target: Linux. Recover Linux Data Arrays after Ransomware Attack, *The Eurasia Proceedings of Science Technology Engineering and Mathematics*, 2022, Vol. 19, pp. 78-86, ISSN: 2602-3199, https://doi.org/10.55549/epstem.1219172

Approbation

Results obtained in the research have been used in the following international, national and university projects:

• Project MU21-FMI-007 "Symbiosis between Mathematics and Computer Science (SMI at FMI)" at the Research Fund of Plovdiv University, university, 2021-2022;

• Project MU21-FMI-009 "Systematic Innovative Research in Mathematics and Computer Science" at the Research Fund of Plovdiv University, university, 2021-2022;

• Project MUPD23-FMI-009 "Development of ICT through New Research and Technological Solutions" at the Research Fund of Plovdiv University, university, 2023-2024;

• Project CP-06 NP62/1 "Mathematical and Information Modeling of Dynamic Processes – New Theoretical Results, Research Methods, and Applications" at the Scientific Research Fund of the Bulgarian Academy of Sciences, national, 2022-2025.

Part of the results obtained in the dissertation thesis have been reported at the following international and national conferences and seminars:

• International Conference on Applied Scientific Computational Intelligence using Data Science (ASCI-2020), India, held online, 22-23 December 2020.

• 47th International Conference Applications of Mathematics in Engineering and Economics (AMEE-2021, Sozopol, 7-13 Jun 2021

• 2nd International Conference on Technology (IConTech), Turkey, Antalya, 16-19 November 2022

• International Scientific Conference Informatics, Mathematics, Education And Their Applications (IMEA'22), Faculty of Mathematics and Informatics, Paisii Hilendarski University of Plovdiv, Bulgaria, Pamporovo,23-25 November 2022

The results obtained in the dissertation thesis have been partially funded under Project No. KP-06 PN 62/1, "Mathematical and Information Modeling of Dynamic Processes - New Theoretical Results, Research Methods, and Applications", 2022-2025.

References

[Alkhalil, 2021] Z. Alkhalil, C. Hewage, L. Nawaf, I. Khan, Phishing attacks: A recent comprehensive study and a new anatomy, *Frontiers in Computer Science 3*, 2021, doi: https://doi.org/10.3389/fcomp.2021.563060

[Altwairqi, 2019] A. Altwairqi, M. AlZain, Ben Soh, M. Masud, J. Al-Amri, Four most famous cyber attacks for financial gains, *International Journal of Engineering and Advanced Technology (IJEAT) Technol 9*, pp. 2131-2139, 2019, ISSN: 2249-8958

[Alwashali, 2021] A. Alwashali, N. Rahman, N. Ismail, A Survey of Ransomware as a Service (RaaS) and Methods to Mitigate the Attack, *2021 14th International Conference on Developments in eSystems Engineering (DeSE)*, Sharjah, United Arab Emirates, 2021, pp. 92-96, doi: 10.1109/DeSE54285.2021.9719456.

[Aurangzeb, 2017] S. Aurangzeb, M. Aleem, M. Iqbal, A. Islam, Ransomware: A Survey and Trends, *Journal of Information Assurance and Security (ESCI – Thomson Reuters Indexed)*, Vol. 12, pp. 48-58, 2017, ISSN: 1554-101

[Gibson, 1997] G. Gibson, D. Nagle, K. Amiri, F. Chang, E. Feinberg, H. Gobioff, C. Lee, File server scaling with network-attached secure disks*, ACM SIGMETRICS international conference on Measurement and modeling of computer systems*, Vol. 25, pp. 272-284, 1997, doi: https://doi.org/10.1145/258623.258696

[Gobioff, 1997] H. Gobioff, G. Gibson, D. Tygar, Security for Network Attached Storage Devices (CMU-CS-97-185), *Carnegie Mellon University, Journal contribution*, 1997, doi: https://doi.org/10.1184/R1/6619784.v1

[Golev, 2022] A. Golev, R. Hristev, M. Veselinova, K. Kolev, Crypto-ransomware Attacks on Linux Servers: A Data Recovery Method, *International Journal of Differential Equations and Applications*, Vol. 21, No. 2 (2022), pages: 19-29, ISSN (Print): 1311-2872; ISSN (Online): 1314-6084, https://www.ijpam.eu/en/index.php/ijdea/article/view/6002/283

[Goyal, 2014] S. Goyal, Public vs Private vs Hybrid vs Community – Cloud Computing: A Critical Review, *International Journal of Computer Network and Information Security*, 6, pp. 20-29, 2014, doi: 10.5815/ijcnis.2014.03.03.

[Grance, 2011] T. Grance, W. Jansen, *Guidelines on Security and Privacy in Public Cloud Computing*, Special Publication (NIST SP), National Institute of Standards and Technology, Gaithersburg, MD, 2011 doi: https://doi.org/10.6028/NIST.SP.800-144

[Hristev, 2021] R. Hristev, M. Veselinova, ICT for Cyber Security in Business, *2021 IOP Conf. Ser.: Mater. Sci. Eng. 1099 012035*, ISSN (Online): 1757-899X, doi:10.1088/1757-899X/1099/1/012035

[Hrıstev, 2022a] R. Hrıstev, M. Veselınova, K. Kolev, Ransomware Target: Linux. Recover Linux Data Arrays after Ransomware Attack, *The Eurasia Proceedings of Science Technology Engineering and Mathematics*, Vol. 19, pp. 78-86, 2022, ISSN: 2602-3199, doi: https://doi.org/10.55549/epstem.1219172

[Hristev, 2022b] R. Hristev, M. Veselinova, Using private cloud for information arrays recovery from ransomware attacks, *AIP Conference Proceedings 2505, 060006 (2022)*, ISSN: 1551-7616, doi: https://doi.org/10.1063/5.0100654

[Hristev, 2023] R. Hristev, M. Veselinova, K. Kolev, Ransomware Attacks on Windows Server: Infection and Recovery, *International Journal of Differential Equations and Applications*, Vol. 22, No. 1 (2023), pp. 57-66, ISSN (Print): 1311-2872; ISSN (Online): 1314-6084, https://www.ijpam.eu/en/index.php/ijdea/article/view/6028/306

[Kaur, 2015] M. Kaur, H. Singh, A Review of Cloud Computing Security Issues, *International Journal of Grid and Distributed Computing*, 8, pp. 215-222, 2015 doi: 10.14257/ijgdc.2015.8.5.21

[Kiru, 2019] M. Kiru, J. Aman, The Age of Ransomware: Understanding Ransomware and Its Countermeasures, *Artificial Intelligence and Security Challenges in Emerging Networks*, IGI Global, 2019, ISBN-13: 9781522573531, doi: 10.4018/978-1-5225-7353-1.ch001

[Knudsen, 2011] L. Knudsen, M. Robshaw, Brute force attacks., *The Block Cipher Companion*, Springer Berlin, Heidelberg, October 2011, ISBN: 978-3-642-17342-4, doi: https://doi.org/10.1007/978-3-642-17342-4

[Odun-Ayo, 2017] I. Odun-Ayo, O. Ajayi, M. Akanle, R. Ahuja, An Overview of Data Storage in Cloud Computing*, International Conference on Next Generation Computing and Information Systems (ICNGCIS)*, pp. 29-34, December 2017, doi: 10.1109/ICNGCIS.2017.9

[Prachi, 2014] P. Deshpande Prachi, S. Sharma, K. Peddoju, Implementation of a Private Cloud: A Case Study, *Advances in Intelligent Systems and Computing*, 259, pp. 635-647, 2014, doi: 10.1007/978-81-322-1768-8_56

[Preston, 2022] W. Preston, *Using SANs and NAS*, O'Reilly Media, 2022, ISBN: 9780596001537

[Richardson, 2017] R. Richardson, M. North, Ransomware: Evolution, mitigation and prevention, *International Management Review*, Vol. 13, No. 1, pp. 10-21, 2017, https://digitalcommons.kennesaw.edu/facpubs/4276

[Rieger, 2011] S. Rieger, H. Richter, Y. Xiang, Introducing Federated WebDAV Access to Cloud Storage Providers, *International Conference on Cloud Computing, GRIDs, and Virtualization (CLOUD COMPUTING'11),* Proceedings 2, pp. 46-51, 2011, ISBN: 978-1-61208-153-3

[Salahdine, 2019] F. Salahdine, N. Kaabouch, Social engineering attacks: A survey, *Future Internet 2019*, 11(4), 89, 2019, doi: https://doi.org/10.3390/fi11040089

[Stiawan, 2017] D. Stiawan, M. Yazid Idris, A. Hanan Abdullah, F. Aljaber, R. Budiarto, Cyber-Attack Penetration Test and Vulnerability Analysis, *International Journal of Online Engineering* Vol. 13 No. 1, 2017, doi: https://doi.org/10.3991/ijoe.v13i01.6407

[Venkatesha, 2021] S. Venkatesha, K. Rahul Reddy, B. Chandavarkar, *Social engineering attacks during the COVID-19 pandemic*, SN Comput Sci. 2021;2(2):78. doi: 10.1007/s42979-020-00443-1. Epub 2021 Feb 6. Erratum in: SN Comput Sci. 2021; 2(3):134. PMID: 33585823; PMCID: PMC7866964

[Vurukonda, 2016] N. Vurukonda, Dr. B. Rao, A Study on Data Storage Security Issues in Cloud Computing, *Procedia Computer Science*, 92, pp. 128-135, 2016, doi: 10.1016/j.procs.2016.07.335

[Whitehead, 1998] E. Whitehead, M. Wiggins, WebDAV: IEFT standard for collaborative authoring on the Web*, in IEEE Internet Computing*, Vol. 2, No. 5, pp. 34-40, Sept.-Oct. 1998, doi: 10.1109/4236.722228.

[Whitehead, 1999] Jr. Whitehead, E. James, Y. Goland, WebDAV: A network protocol for remote collaborative authoring on the Web, *ECSCW'99: Proceedings of the Sixth European Conference on Computer Supported Cooperative Work*, 12–16 September 1999, Copenhagen, Denmark, pp. 291-310, Springer Netherlands

[Божиков, 2019] А. Божиков, *ВЪЗСТАНОВЯВАНЕ НА ИНФОРМАЦИОННАТА ИНФРАСТРУКТУРА ПРИ БЕДСТВИЯ И АВАРИИ*, СТОПАНСКА АКАДЕМИЯ „Д. А. ЦЕНОВ" – СВИЩОВ, КАТЕДРА „БИЗНЕС ИНФОРМАТИКА", 2019

[Денчев, 2019] С. Денчев, *Информация и сигурност, Университет по библиотекознание и информационни технологии*, Академично издателство „За буквите – О писменехь", 2019, ISBN: 978-619-185-369-4

[Младенова, 2011] М. Младенова, *ОБЛАЧНИ ИЗЧИСЛЕНИЯ (CLOUD COMPUTING): СЪЩНОСТ, ПРЕДИМСТВА, НЕДОСТАТЪЦИ И РИСКОВЕ, СЪСТОЯНИЕ И ПЕРСПЕКТИВИ*, Интел Ентранс, 2011, ISBN: 978-954-2910-07-7

[CVE, 2023] CVE – Search Results – VPN https://cve.mitre.org/cgi-bin/cvekey.cgi?keyword=vpn, last accessed October 2023

[Nextcloud, Encryption Configuriation, 2023] Encryption configuration — Nextcloud latest Administration Manual latest documentation, https://docs.nextcloud.com/server/latest/admin_manual/configuration_files/encryption_configuration.html, last accessed October 2023

[Nextcloud, File Versions, 2023] Controlling file versions and aging — Nextcloud latest Administration Manual latest documentation, https://docs.nextcloud.com/server/latest/admin_manual/configuration_files/file_versioning.html, last accessed October 2023

[Nextcloud, OCC command, 2023] Using the occ command — Nextcloud latest Administration Manual latest documentation, https://docs.nextcloud.com/server/latest/admin_manual/configuration_server/occ_command.html, last accessed October 2023