ПЛОВДИВСКИ УНИВЕРСИТЕТ „ПАИСИЙ ХИЛЕНДАРСКИ“
ФАКУЛТЕТ ПО МАТЕМАТИКА И ИНФОРМАТИКА
КАТЕДРА „КОМПЮТЪРНИ ТЕХНОЛОГИИ“

РОСЕН ПЕТРОВ ХРИСТЕВ

ВЪЗСТАНОВЯВАНЕ НА ИНФОРМАЦИОННИ МАСИВИ В ОБЛАЧНА СРЕДА

А В Т О Р Е Ф Е Р А Т

на дисертационен труд
за присъждане на образователната и научна степен „доктор“
по област на висше образование: 4. Природни науки, математика и информатика
професионално направление: 4.6. Информатика и компютърни науки,
докторска програма: Информатика

Научен ръководител: проф. д-р Ангел Голев

Пловдив, 2023 г.

Дисертационният труд е обсъден и насочен за защита на разширено заседание на катедра „Компютърни технологии“ при Факултета по математика и информатика на ПУ „Паисий Хилендарски“.

Дисертационният труд съдържа 113 страници. Използваната литература включва 72 източника – 55 заглавия на латиница, 3 на кирилица и 14 интернет адреса.

Списъкът на авторските публикации се състои от 3 заглавия.

Защитата на дисертационния труд ще се състои на 22.01.2024 г. от 11:00 в Заседателната зала на новата сграда на ПУ „Паисий Хилендарски“, гр. Пловдив.

Материалите по защитата са на разположение на интересуващите се в секретариата на ФМИ, нова сграда на ПУ, каб. 330, всеки работен ден от 8:30 до 17:00 часа.

Автор: Росен Петров Христев

Заглавие: „Възстановяване на информационни масиви в облачна среда“

Университетско издателство „Паисий Хилендарски“

Пловдив, 2023 г.

Съдържание

Обща характеристика на дисертационния труд 5
Актуалност на проблема 5
Цел и задачи на дисертационния труд 5
Структура и обем на дисертационния труд 6
Кратко съдържание на дисертационния труд 7
Глава I. Съхранение на данните и заплахи за тяхната сигурност в ИТ инфраструктури 7
Методи за съхранение в локална компютърна мрежа 7
Облачно съхранение на данни 8
Стандартни заплахи за данните, циркулиращи в една ИТ инфраструктура 10
Зловреден софтуер 10
Резултати и изводи 12
Глава II. Достъпване на данни, съхранявани в ИТ инфраструктури 12
Достъпване на данни, съхранявани при стандартни инфраструктури 12
Достъпване на данни, съхранявани в публични и частни облаци 13
Разлики при съхранение и достъпване на данни в стандартни инфраструктури и в публични и частни облаци 14
Авторски метод за внедряване на облак във вече съществуваща ИТ инфраструктура (MICI) 16
Резултати и изводи 19
Глава III. Възстановяване на изтрити и презаписани данни в ИТ инфраструктура 20
Възстановяване на данни при стандартни инфраструктури 20
Възстановяване на изтрити данни при облачни инфраструктури 21
Възстановяване на презаписани данни 24
Резултати и изводи 26
Заключение 27
Перспективи за развитие 28
Основни приноси на дисертационния труд 28
Публикации по дисертационния труд 29
Апробация 29
Библиография 30

Списък на съкращенията

AFP – Apple Filing Protocol (мрежов протокол за споделяне на файлове)

HTTP – Hypertext Transfer Protocol (протокол за пренос на хипертекст)

LDAP – Lightweight Directory Access Protocol (протокол за достъп до онлайн справочни услуги)

MICI – Method for Implementing Cloud in existing Infrastructure

NAS – Network-attached storage (мрежово прикачено хранилище)

NFS – Network File System (мрежов протокол за споделяне на файлове)

OCC – ownCloud Console (конзолен интерфейс за ownCloud и NextCloud)

RaaS – Ransomware as a Service (модел за разпространение на криптовируси)

RDSC – Recovery Data Stored in Cloud

RVSC – Recovery Versions Stored in Cloud

SaaS – Software as a Service (модел за доставка на софтуер)

SAN – Storage Area Network (мрежа за съхранение на данни)

SMB – Server Message Block (мрежов протокол за споделяне на ресурси)

VPN – Virtual Private Network (виртуална частна мрежа)

WebDAV – Web Distributed Authoring and Versioning (мрежови протокол, работещ върху HTTP протокол)

СУБД – Система за Управление на База от Данни

Обща характеристика на дисертационния труд

Актуалност на проблема

Развитието на корпорациите води до нуждата от внедряване на различни информационни технологии с цел оптимизиране на операциите. Процесът по дигитализация, който беше значително ускорен с обявяването на пандемията от Covid-19 през 2020 година, прави сървърите за съхранение на информационни масиви едни от най-разпространените. Това улеснява данните, циркулиращи в една ИТ инфраструктура, за да могат да бъдат лесно споделени между служителите на компанията.

Такъв тип сървъри намират приложение както във фирми с по няколко служители, така и в големи корпорации. Обикновено данните, съхранявани в инфраструктурата на една фирма, надвишават многократно финансовата стойност на хардуера, необходим за тяхното съхранение.

Данните, от своя страна, са сред уязвимите ресурси на ИТ инфраструктурата. Освен стандартните заплахи над тях, свързани с природни бедствия, човешки грешки, дефектирал хардуер, софтуерни грешки и други, допълнително се развиват и разработчиците на зловреден софтуер. През последните години атаките, насочени към информационни масиви, стават все по-голям обем. Една от най-големите атаки с криптовируса WannaCry през месец май 2017-та година успява да порази над 200 хиляди компютри в над 150 държави по света. След навлизането на RaaS на пазара в края на 2020 година се наблюдава повишаване на атаки с криптовируси с над 50% в световен мащаб.

Стандартните методи за защита, например бекъп на информацията, не винаги са достатъчно сигурни. Бекъпите също могат да бъдат криптирани, ако не са управлявани правилно. Допълнително това крие риска информацията, която е от времето на последното архивиране до момента на заразяване, също да бъде загубена.

Цел и задачи на дисертационния труд

Основната цел на дисертационния труд е (А) да се създаде метод за внедряване на облак в съществуваща инфраструктура, както и (Б) да се разработят методи и инструменти за възстановяване на изтрити информационни масиви и предходни версии на файлове, съхранявани в облачна инфраструктура.

Основните задачи са:

1. Изследване и класификация на видовете заплахи за данните, съхранявани в ИТ инфраструктура. Подзадачи:
   1.1. Изследване на стандартните заплахи, водещи до загуба на информационни масиви;
   1.2. Изследване на методите за разпространение на зловреден софтуер;
   1.3. Класификация на зловредния софтуер, водещ до загуба на данни;
2. Проучване на методите и подходите за съхранение и обработка на информационни масиви в ИТ инфраструктурите. Подзадачи:
   2.1. Изследване на стандартните и облачните инфраструктури за съхранение и обработка на информационни масиви в ИТ инфраструктурите;

2.2. Изследване на възможностите за скалиране на облачни инфраструктури;
2.3. Сравнителен анализ на различните методи за съхранение и обработка на информационни масиви;

3. Анализиране на възможностите за възстановяване на изтрити данни от ИТ инфраструктура. Подзадачи:
   3.1. Анализиране на разликите при възстановяване на изтрити файлове, съхранявани в стандартни и облачни инфраструктури;
   3.2. Апробация на метод за възстановяване на изтрити данни съхранявани в облачна инфраструктура;
   3.3. Проектиране на скриптове за автоматизирано възстановяване на изтрити данни съхранявани в облачна инфраструктура;
4. Анализиране на възможностите за възстановяване на предходни версии на файлове. Подзадачи:
   4.1. Апробиране на метод за възстановяване на предходни версии на файлове, съхранявани в облачна инфраструктура;
   4.2. Проектиране на скрпит за автоматизирано възстановяване на предходни версии на файлове, съхранявани в облачна инфраструктура.

Структура и обем на дисертационния труд

Дисертационният труд се състои от увод, три глави, заключение, списък на авторските публикации по темата и апробация на резултатите, библиография, и е в обем от 113 страници. Към работата са добавени авторска справка за приносите на дисертационния труд, декларация за оригиналност и перспективи за бъдещо развитие.

В **Глава 1** е направено проучване и анализ на методите за съхранение на данни в една съвременна ИТ инфраструктура. Разгледани са разликите между публичните и частни облаци. Анализирани са най-честите заплахи при информационните масиви, които водят до загуба на данни.

В **Глава 2** са изследвани методите за достъпване на информационните масиви, циркулиращи в ИТ инфраструктурата. Анализирани са основните разлики при съхранение и достъпване на данни, съхранявани в стандартни и облачни инфраструктури. Разработен е авторски метод MICI за внедряване на облачна инфраструктура във вече съществуваща инфраструктура и преминаване от стандартна инфраструктура към облачно съхранение на данните. Методът е апробиран в повече от 10 реални ИТ инфраструктури, които са различни по големина и сфери на дейност. Част от организациите, при които е внедрен частен облак и е приложен методът, са: Факултет по математика и информатика към ПУ „Паисий Хилендарски“, УМБАЛ „Каспела“, Мебелна фабрика „Колорадо“, печатница „РайтПак“ ЕООД.

В **Глава 3** са разгледани възможностите за възстановяване на изтрити и презаписани информационни масиви, записани в облачни инфраструктури. В този контекст, бяха разработени два авторски метода за възстановяване на данни, циркулиращи в облачна среда: RDSC и RVSC. RDSC е метод, специално създаден за възстановяване на изтрити данни, които са съхранявани в облачна среда. За да се

постигне тази цел, са създадени скриптове, които автоматизират процеса на възстановяване на изтрити данни, съхранявани на частни облаци, използвайки метода RDSC. RVSC е вторият авторски метод, който е разработен и описана в Глава 3. Той е предназначен за възстановяване на предходни версии на данни, съхранявани в облачна среда. Създадени са авторски скриптове за автоматизация на задачата. Тези методи и скриптове представляват важни инструменти за възстановяване на данни в облачни среди и допринасят за по-голямата сигурност и устойчивост на данните, съхранявани в облачни инфраструктури.

Кратко съдържание на дисертационния труд

Глава I. Съхранение на данните и заплахи за тяхната сигурност в ИТ инфраструктури

Методи за съхранение в локална компютърна мрежа

Съхранението на данни в локална компютърна мрежа може да бъде осъществено чрез различни методи и технологии в зависимост от специфичните нужди и изисквания на корпорацията. Изборът на подобен метод за съхранение е предопределен от различни фактори, между които са специфичните нужди на организацията и потребителите, бюджета, сигурността и обема на данните. В повечето случаи, съчетаването на различни методи може да бъде най-подходящото решение, осигуряващо възможност за съхранение, сигурност и лесен достъп до данните.

Споделени директории - обикновено се разглеждат като отделно хранилище за информация с цел споделяне. Те представляват метод за съхранение на данни в локална компютърна мрежа, където файлове и директории могат да бъдат споделени и достъпни от различни компютри в съответната мрежа. За достъп до споделените директории в локалната компютърна мрежа се използват мрежови протоколи, които позволяват на компютрите да комуникират и споделят данни. Мрежовият протокол зависи от операционната система, нуждите, за които ще се прилага, и средата на мрежата. Различните протоколи имат индивидуални характеристики и съвместимости. Някои от тези протоколи са следните: *Server Message Block* (SMB), *Network File System* (NFS), *Apple Filing Protocol* (AFP) и други.

Мрежово прикачено хранилище (NAS) и Мрежа за съхранение на данни (SAN) са два различни метода за съхранение на данни, които се имплементират в информационните технологии и компютърните мрежи. Те имат различни предназначения и характеристики, и се прилагат в зависимост от конкретните нужди и изисквания на организациите.

Мрежово прикачено хранилище (NAS) е устройство, което се използва за съхранение, управление и споделяне на файлове и данни в локална компютърна мрежа. Обикновено е сървър, който се хоства в мрежата и предоставя файлови услуги на потребителите чрез множество мрежови протоколи. Той представлява специализирано устройство, проектирано за съхранение на данни и файлове.

Мрежа за съхранение на данни (Storage Area Network или SAN) е тип мрежова архитектура, създадена за централизирано съхранение и управление на данни. Тя

осигурява висококапацитетно и високоскоростно съхранение, където данните се съхраняват на отделни устройства, наречени сървъри за съхранение, и могат да бъдат споделяни от множество сървъри в мрежата. SAN се имплементира често в големи организации, които имат големи обеми данни, при които е необходима висока производителност и надеждност на съхранението. Този тип мрежа е особено подходяща за сървъри, при които е внедрена виртуализация и база от данни.

Облачно съхранение на данни

Облачното съхранение на данни (Cloud Storage) е метод за съхранение на данни на отдалечени сървъри, които обикновено се хостват в специализирани центрове за данни и са достъпни през интернет. Този метод позволява на потребителите да съхраняват, управляват и споделят данни, като използват ресурси, предоставени от доставчика на облачни услуги.

Видове облаци - Облачните изчисления са революционен механизъм, който промени начина за закупуване на софтуер и проектиране на софтуер от предприятията. Тези нови методи променят из основи начина за съхраняване на данни. В наши дни почти всичко се съхранява в „облак“, вместо да се изпълняват програми и да се съхраняват данни на определен компютър. Облачните технологии предоставят множество предимства на клиентите си като безплатни услуги, лесен достъп през интернет, скалируемост на наличните ресурси и други [Vurukonda, 2016].

Различаваме четири основни типа облаци: публични, частни, общностни и хибридни облаци. Управлението на публичната облачна инфраструктура се извършва от доставчика, който я предлага за обществено ползване [Kaur, 2015]. Частната облачна инфраструктура е достъпна само от членове на организацията, може да бъде интегрирана при трети страни само срещу предоставен достъп [Prachi, 2014]. Хибридният облак е инфраструктура, която представлява композиция от две или повече различни облачни инфраструктури (частни, публични), имащи уникални обекти, но са свързани помежду си със стандартни или собствени технологии, позволяващи пренасянето на данни или приложения между компонентите [Младенова, 2011]. Общностният облак донякъде е подобен на частния облак, но инфраструктурата и изчислителните ресурси се споделят от няколко организации, които имат общи съображения относно поверителността, сигурността и регулаторните изисквания [Goyal, 2014].

Публични облаци – те са типа SaaS (Software as a service) услуги – облачни модели за доставка на софтуер, разрешаващ на крайните потребители достъп до софтуерни приложения през интернет. Най-разпространените публични облаци са Google Drive, Microsoft OneDrive, DropBox, Amazon Cloud, Apple iCloud и други. Този тип облаци са добри за съхранение и споделяне на нечувствителни данни, но не са препоръчителен вариант за съхранение и споделяне на чувствителна и конфиденциална информация поради липсата на контрол върху сървърите, на които се съхраняват данните.

Частни облаци - Съществуват различни доставчици, които предлагат решения за съхранение на файлове. Разгледаните частни облаци в дисертационния труд могат да бъдат интегрирани в ИТ инфраструктура безплатно, без външна организация да има

достъп до файловете, циркулиращи в инфраструктурата. Част от най големите предимства на частните облаци са: контрол, сигурност и криптиране, достъп до данните, интеграция и други.

Структура на файловата система на частни облаци - разгледаните частни облаци запазват данни и файлове във файлова система на сървъра и СУБД [Odun-Ayo, 2017]. По този начин файловете на сървъра се съхраняват директно във файловата система на сървъра, а мета данните, също както и данните за самите файлове, публичните споделяния и споделянията между потребителите и други – в базата от данни. За всеки потребител е възможно да съществуват следните директории:

- files – директория за всички ресурси, които са собственост на потребителя. Ако към даден потребител бъде споделена директория, то собственик на ресурса остава споделящия. В случай че бъде създаден файл или директория в споделен ресурс, то собствеността ще бъде на този, който е първият собственик на ресурса.
- files_trashbin – директорията за ресурси, които са изтрити от потребителя. Важно е да се отбележи, че ако потребител изтрие ресурс, който е споделен с него, то изтритите данни няма да се намират в профила на собственика на ресурса, а в директорията на потребителя, който е изтрил ресурсите. Изтритите файлове и директории за частния облак се изтриват автоматично и от сървъра, ако са били повече от 30 дни във files_trashbin или ако потребителят изчерпа повече от 50% от свободната си квота. Ресурсите се изтриват по метода на опашката. Директорията е възможно да има до 3 поддиректории: files, keys и versions.
- file_versions – в директорията се запазват версии за всички файлове, които са собственост на потребителя. Всяка версия завършва с 12 символен суфикс, който включва Unix Timestamp на времето на създаването на версията. Версията е пълно копие на първоизточника на файла. Ресурсите нямат ограничение за това колко версии да имат, като те не могат да надхвърлят повече от 50% от свободното пространство от квотата на потребителя. При намаляване на свободното пространство на потребителя най-старите версии се изтриват по метода на опашката.
- files_encryption – директорията е налична само ако е включен модулът за криптирано съхранение на данни на сървъра. При активиране той няма да криптира настоящите файлове, които са налични в частния облак, а ще криптира всички файлове, които ще бъдат качени от момента на активиране на функцията. Директорията съдържа 2 поддиректории: keys и OC_DEFAULT_MODULE.
- uploads – директорията се използва за временно съхраняване на файлове. Когато потребител качи нов ресурс в облака, то частите му първоначално ще се намират в uploads. Едва когато качването на ресурса приключи, ресурсът ще бъде позициониран в правилната директория във files. Впоследствие ресурсът ще бъде автоматично изтрит от директорията uploads.

Разбирането на файловата структура на частния облак би спомогнало и улеснило поставените задачи за възстановяване на изтрити и презаписани файлове, също както и връщане към предходни версии на файлове.

Стандартни заплахи за данните, циркулиращи в една ИТ инфраструктура

Данните, съхранявани в една ИТ инфраструктура, са едни от най-ценните ресурси, като едновременно с това са и сред ресурсите, за които има най-много заплахи. Финансовата стойност на ИТ инфраструктурата обикновено многократно надхвърля стойността, необходима за съхранението и обработването на данни. Данните, от своя страна, са сред уязвимите ресурси на инфраструктурата, като стандартните заплахи над тях са свързани с природни бедствия, човешки грешки, дефектирал хардуер, софтуерни грешки, проблеми с електрозахранването и други [Божиков, 2019].

Зловреден софтуер

Зловреден софтуер е всеки софтуер, който е създаден с цел нанасяне на определени щети върху дадена компютърна система. Видовете му могат да варират – от кражба на информация чрез нейното модифициране до нейното цялостно разрушение [Денчев, 2019].

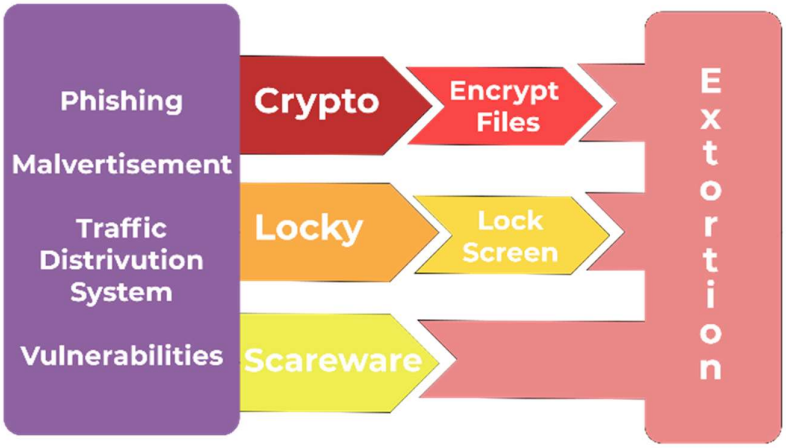
Съществуват различни видове атаки, чрез които се разпространява зловредният софтуер. Сред най честите са:

- Фишинг (phishing attack) – това е най-разпространената по вид атака и представлява изпращане на електронни съобщения. Възползват се най-вече от познати източници като податели на имейлите, с цел да се заблуди потребителят и да се получи достъп до чувствителна информация [Alkhalil, 2021].
- Използване на чужда самоличност – атакуващият субект се представя за друг човек. Този тип атаки са по-успешни тогава, когато са предназначени за конкретна жертва и тя е изследвана. По този начин е по-лесно да бъде заблудена, като се изпратят привидно достоверно изглеждащи писма от потребител, с който жертвата е в редовна комуникация [Altwairqi, 2019].
- Brute Force – този тип атаки се прилагат при лошо защитени мрежи. Съществуват множество ботове, които сканират постоянно публичните адреси за наличие на пуснати услуги. Най-често се сканират стандартните портове, описани в IANA, и при наличие на отворен порт се правят множество опити за влизане с речници за пароли и потребители [Knudsen, 2011].
- Социално инженерство – атаките от този тип залагат на манипулация на потребителя. Злонамерените действия започват със събиране на информация за жертвата от корпоративни сайтове и публикации, или дори комуникация с жертвата. След като има достатъчно събрана информация и е направена стратегия за атаката се прибягва към самата атака. Тя често се възползва от набелязани слабости в системата или манипулиране на потребител, който има достъп до инфраструктурата [Salahdine, 2019].

- Проникване в средата – атаките от този тип обикновено се характеризират с търсене на различни видове уязвимости в системата, за да бъде получен неоторизиран достъп [Stiawan, 2017].

Компютърните вируси покриват широка гама от вреди, която зависи основно от вида на атакуващия вирус. Криптовирусите като фамилия от вируси безспорно вредят най-много на файловете в наши дни. Заразяването с тях обикновено е чрез изпълнение на файлове, съдържащи зловреден код. Основната им цел е шифроване на данни или устройства и изискване на откуп, за да бъде възстановен достъпът до данните [Tailor, 2017].

Настъпването на пандемията от Covid-19 бе предпоставка множество фирми да понижат нивото на сигурност на ИТ инфраструктурите си, за да могат служителите бързо да преминат в дистанционен режим на работа, с което се забеляза нарастване на атаките с криптовируси [Venkatesha, 2021]. Една от основните причини за това е и наличието на нов вид услуга – Ransomware as a Service (RaaS), което се наложи в края на 2020 година [Alwashali, 2021]. Вследствие на това се наблюдава повишаването на криптовирусните атаки с над 50% в световен мащаб.



Фигура 1 Видове криптовируси

Crypto-ransomware - типът crypto-ransomware се счита за много опасен, тъй като криптира файловете на жертвата, което прави невъзможен достъпа до тях без валиден ключ за декриптиране [Aurangzeb, 2017].

Locker-ransomware - заключва системата на жертвата и показва страница за вход. Жертвата трябва да плати откуп, за да получи парола, с която да отключи системата [Richardson, 2017]. Locker-ransomware се счита за не толкова опасен, тъй като атаката често може да бъде разрешена чрез рестартиране на системата в безопасен режим.

Scareware - този рансъмуеър не представлява реална опасност за жертвата си. Основната му функция е да изплаши жертвата си, която следва да плати откуп [Hristev, 2022a]. Scareware атака често се стартира чрез изскачащи прозорци, които се появяват на екрана на потребителя, като го предупреждават, че компютърът или файловете му са заразени и след това предлагат „решение на проблема“ [Kiru, 2019].

Резултати и изводи

Сървърите за съхранение на информационни масиви са едни от най-разпространените в ИТ инфраструктурите. Благодарение на тях, данните, циркулиращи в инфраструктурата, могат да бъдат лесно споделяни между служителите на компанията.

От направения обзор и анализ на системите за съхранение на данни могат да бъдат синтезирани следните изводи:

- Стандартните методи за съхранение на данни като споделени директории, NAS и SAN сървъри не са достатъчно гъвкави.
- Публичните облаци са добро решение за съхраняване на неконфиденциална информация.
- Наличното пространство за съхранение при публични облаци може да бъде лесно скалирано.
- Частните облаци са добро решение за съхранение на конфиденциални и фирмени данни.
- Недостатък на частните облаци може да бъде по-трудното им скалиране и необходимостта от технически подготвени лица за тяхната поддръжка.
- Част от най-големите хакерски атаки, водещи до загуба на данни, са реализирани посредством криптовируси.

В Глава 1 са постигнати следните резултати:

1. Проучени са методите за съхранение на информационни масиви в една ИТ инфраструктура.
2. Разгледани и сравнени са част от най-разпространените доставчици на публични облачни услуги от типа SaaS и частни облаци.
3. Изследвани са възможностите за скалиране на частни облаци и файловата структура, в която те съхраняват и обработват данните.
4. Анализирани са обичайните заплахи, водещи до загуба на данни в ИТ инфраструктури.
5. Изследвани са стандартните методи за разпространение на зловреден софтуер и са класифицирани три вида изнудващ зловреден софтуер.

Глава II. Достъпване на данни, съхранявани в ИТ инфраструктури

В Глава 2 от дисертационния труд се разглеждат различни методи за достъпване на данните, които циркулират в една съвременна ИТ инфраструктура. Направено е сравнение между стандартните методи за съхранение на данни и съхранението на данни при облачни инфраструктури. Разгледан и описан е авторски метод, чрез който е възможно интегриране на облак за съхранение на информационни масиви в съществуваща инфраструктура.

Достъпване на данни, съхранявани при стандартни инфраструктури

Стандартните методи за съхранение на данни се ограничават до разпространение на данните в една локална компютърна мрежа. Това води до редица ограничения, които затрудниха множество организации с обявяването на пандемията от Covid-19 в

началото на 2020 година. Преминаването към дистанционен режим ускори многократно дигитализацията на работата, което, от своя страна, провокира множество корпорации в световен мащаб много бързо да реорганизират установения работен режим. Това накара ИТ специалистите да занижат критериите за ИТ сигурността. В същото време разработчиците на зловреден софтуер стават все по-изобретателни, за да могат да реализират своите цели. Така например множество зловредни атаки, в това число и криповируси, започнаха да използват освен стандартните фишинг атаки, допълнително сканирания за уязвимости на системите. Ясен пример за това е, че само за 2022 година дялът на фирмите, внедрили VPN за достъп до частните мрежи, е нараснал с 2%, като е достигнал до 95% общо за организациите в световен мащаб. Справки за zero day exploits в CVE показват, че от 1999 година до сега са заявени близо 700 уязвимости, свързани с VPN, от които около 38% са обявени от започването на Covid-19 пандемията [CVE, 2023]. Сходно е положението и с един от най-разпространените протоколи, що се отнася до споделяне на мрежови ресурси – 256 уязвимости от 1999, от които 58 са анонсирани от началото на 2020 до момента.

При споделени директории и NAS сървъри, данните се съхраняват на файлово ниво и се достъпват най-често през протоколи като Samba и NFS [Gibson, 1997].

Един от основните проблеми е, че при съхраняването на данни чрез споделени директории, NAS и SAN при необходимост от възстановяване на информационни масиви се разчита основно на резервни копия и архиви на данните [Preston, 2022]. Повече от 70% от малкия бизнес, които са основни ползватели на споделени директории и NAS, нямат разписани планове, стратегия и процедура за възстановяване след загуба на данни. Дори при разписани планове за запазване на архивни копия, то почти никога не се правят периодични проверки дали консистентността на данните в резервните копия е налична и до колко е възможно възстановяването на данни от резервните копия.

Данните, съхранявани в споделени директории и NAS устройства, най-често се достъпват като монтирани дискове през мрежата [Gobioff, 1997].

Достъпване на данни, съхранявани в публични и частни облаци

Рядко се срещат публични облаци, които не поддържат достъп до файловете чрез уеб базирано приложение. Този вид достъпване на файловете не винаги е удобно и удачно, тъй като би създало допълнителни стъпки за потребителя при използване на файл, съхранен в облачното хранилище [Grance, 2011]. Освен достъпване на данните чрез уеб браузър, някои от доставчиците на публични облаци предоставят и синхронизиращи софтуери. Благодарение на тях данните, които се съхраняват в облака, се синхронизират с директория на локалната машина на потребителя. Това изтегля копие на файла на работната станция на потребителя. Когато потребителят направи промяна във файл или създаде нов, то синхронизиращият клиент ще качи промените, като при повечето доставчици старите данни ще бъдат запазени като различна, предходна версия на файла.

Разгледаните частни облаци предлагат три основни варианта, чрез които данните, съхранявани на тях, могат да бъдат достъпени: през уеб браузър, през синхронизиращ клиент и чрез използването на WebDav.

WebDAV (Web Distributed Authoring and Versioning) като разширение е анонсиран през 1996-та година от Джим Лаитхед и всъщност е надграждане на HTTP протокола и се базира на него, но позволява на потребителите да манипулират информационни масиви директно на сървъра [Whitehead, 1999]. Протоколът се имплементира от разгледаните частни облаци и посредством него могат да се монтират дискове в операционната система, които потребителите виждат и да работят с тях както при файловите сървъри [Rieger, 2011].

При Linux и MacOS базирани операционни системи протоколът работи стандартно и може да се употребява без проблеми. При част от Windows базираните операционни системи се срещат проблеми, като при някои по стари версии включително трябва да се извършат допълнителни промени. Така например, при Windows 7 и Windows 8 е необходима редакция включително в регистрите на операционната система. Windows базираните операционни системи могат да комуникират през WebDav протокола чрез стандартен сървис, наречен WebClient, позволяващ създаване и редактиране на интернет базирани файлове [Whitehead, 1998].

Разлики при съхранение и достъпване на данни в стандартни инфраструктури и в публични и частни облаци

Данните, циркулиращи в стандартните инфраструктури, се достъпват основно през протоколи като Samba и NFS. Част от производителите на NAS оборудване през последните години залагат освен на двата основни протокола и синхронизиращи клиенти, и на протоколи, чрез които клъстери за виртуализация и други подобни софтуери също да имат достъп до NAS устройствата.

От друга страна, публичните облаци се достъпват основно чрез уеб браузър или синхронизиращ клиент. Това може да предизвика затруднения на потребителите при определени условия. Например добавките, които са имплементирани за онлайн редактиране на електронни офис документи, не поддържат пълни функционалности. Допълнително ако файл бъде свален, за да бъде обработен с офлайн версия на офис пакет, може да се стигне до ситуация с презаписване на информация. Данните в този случай могат да бъдат възстановени чрез сравняване на предходните версии на файла. Това изисква ръчна намеса на потребител, което също може да причини неконсистентност на данните, съхранявани във файловете.

Частните облаци поддържат и двете функции за работа с данни – уеб базиран интерфейс и синхронизиращ клиент, но за разлика от повечето SaaS доставчици на публични облаци, предлагат и възможност за работа с файловете чрез WebDav. Достъпването на облака през WebDav понякога може да причини затруднения, които са подробно дискутирани в подпараграф 2.2 на дисертационния труд. При правилно избиране на инфраструктурата на частния облак е възможно да се надгради с хибридно решение, при което да се предложат на потребителя удобствата както при работа със стандартни методи за съхранение на информация, така и сигурността от това данните да бъдат съхранявани в частен облак.

Таблица 1 Разлики при достъпване на данни в стандартни инфраструктури и облачна среда

Технология	**Инфраструктура**	**Основни начини за достъпване на данните**	**Предимства**	**Недостатъци**
Споделени директории	Файлова	Samba, NFS	Лесна работа с файловете	• Силно намаляване на сигурността, ако данните не циркулират в частна мрежа • Данните не са защитени
NAS	Файлова	Samba, NFS, Синхронизиращ клиент	Лесна работа с файловете	• Силно намаляване на сигурността, ако данните не циркулират в частна мрежа • Данните не са защитени
SAN	Блокова	Samba, NFS, FiberChannel, iSCSI	Лесна работа с файловете	• Силно намаляване на сигурността, ако данните не циркулират в частна мрежа • Данните не са защитени
Публичен облак	Облачна	Уеб браузър, синхронизиращ клиент	Усложнена структура	• Данните са защитени с двойно изтриване и версии на файловете • Данните могат да бъдат достъпени от публични мрежи, без това да пречи на сигурността им
Частен облак	Облачна	Уеб браузър, синхронизиращ клиент, WebDav	Усложнена структура	• Данните са защитени с двойно изтриване и версии на файловете • Данните могат да бъдат достъпени от публични мрежи, без това да пречи на сигурността им

Авторски метод за внедряване на облак във вече съществуваща ИТ инфраструктура (MICI)

За да бъде внедрен облак във вече съществуваща ИТ инфраструктура, по метода MICI (Method for Implementing Cloud in existing Infrastructure), то първоначално трябва да бъде избран видът облак, който да се интегрира в инфраструктурата. Това може да бъде направено според спецификата на данните, които се очаква да бъдат съхранявани в облачната инфраструктура.

Облачните услуги предоставят възможност за възстановяване на файлове чрез интегрирани функционалности като двойно изтриване на файлове и предходни версии на файлове. За да е възможно възстановяването на информационни масиви, трябва предварително да се изчисли наличното пространство в инфраструктурата. Необходимото пространство на сървъра може да бъде изчислено по следната формула:

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

където: DS – дисково пространство; D – данни; Cr – коефициент увеличения размер на данните, вследствие на криптиращия алгоритъм в зависимост от вида на криптовируса, където Cr > 1.0.

Разгледаните частни облаци лесно могат да бъдат надградени, за да покриват стандарт ISO 27001 – Системи за определяне на сигурността на информацията. В тази ситуация информационните масиви на сървъра ще бъдат съхранявани в криптиран вид, което допълнително би увеличило размера на данни в некриптиран вид с до 35% [Nextcloud, Encryption Configuriation, 2023]. За изчисляване на необходимото дисково пространство за съхранение на данните в криптиран вид, формулата по горе би приела вида:

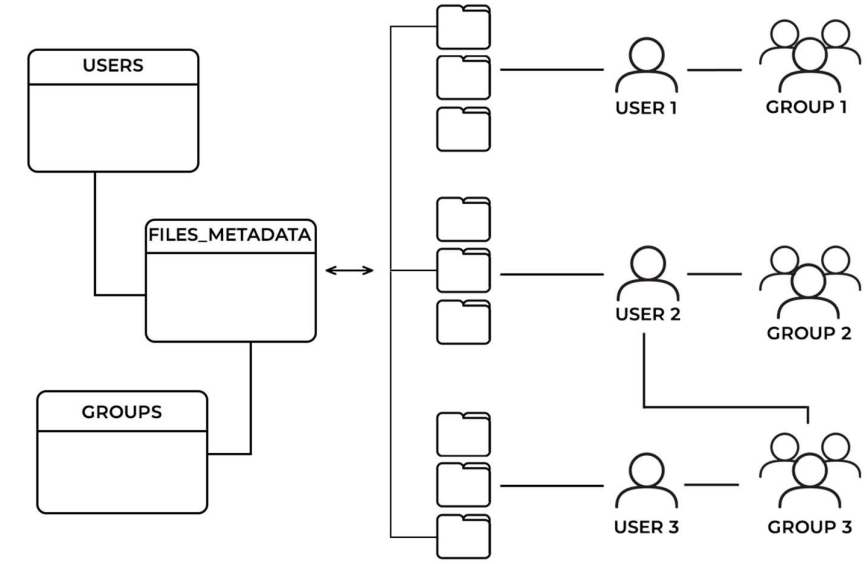
$$DS= D*(Ea+Ea*Cr),$$

където: DS – дисково пространство; D – данни; Ea – коефициент на увеличен размер на файловете, вследствие на активиране на модул за криптирано съхранение на данните, където Ea > 1.0; Cr – коефициент увеличения размер на данните, вследствие на криптиращия алгоритъм в зависимост от вида на криптовируса, където Cr > 1.0.

Структурите, по които облачните системи съхраняват данните си, се различават от файловите структури при стандартните методи за съхранение на данни. Те притежават структура от поддиректории, които не са директно видими за потребителите, като тя е обвързана и със СУБД, използвана от облака. Директорийното дърво може да се различава в зависимост от облака, но на файловата система на сървъра съществува директория за всеки потребител с достъп до облачната инфраструктура. Директорията обикновено е именувана с името на потребителя или уникален идентификатор при LDAP удостоверяване. В дървото на потребителя съществуват директории с наименования като:

- files – място, на което се съхраняват всички файлове, собственост на потребителя;
- files_trashbin – всички изтрити от потребителя файлове;

- file_versions – директория, в която се съхраняват всички различни версии на файлове, създадени от конкретен потребител.



Фигура 2 Файлова структура и база от данни при облачно съхранение на данни

За основната директория на потребителя обикновено може да бъде зададена квота. Съществува разлика между квотите в публичните и частните облаци. При публичните облаци може да се приеме, че квотата на даден потребител всъщност са файловете, които той е създал или съхранява в частния облак, като няма значение дали това е в споделено пространство от друг потребител или е в собствена директория. Когато потребител изтрие ресурс при публичните облаци, ако той е бил споделен с него, ще бъде премахнато споделянето и файлът ще може да се открие в диска на потребителя, собственик на ресурса. Пространството за съответния ресурс вече е използвано от квотата за потребителя, като тази операция няма да намали допълнително наличното свободно пространство на потребителя [Hristev, 2021].

От друга гледна точка, при частните облаци, ако потребител качи ресурс в споделена директория, той не изразходва от собствената си квота, а от тази на потребителя, който е собственик на споделената директория. На практика, собствеността на ресурса се присвоява на потребителя, собственик на споделената директория. Метаданните (например временни файлове, кешове и криптиращи ключове) заемат от квотите на двата потребителя, като по подразбиране е указано, че те не могат да бъдат повече от 10% от съответната квота.

При квотите в частните облаци е препоръчително създаването на основни потребители за системата, на които се съхраняват всички файлове и от тях биват споделени с останалите потребители. Целта на това е при работа с частния облак винаги собствеността на файловете да бъде на основните потребители. По този начин ще бъде позволено при напускане на служител на компанията, неговият потребител лесно да бъде изтрит от системата без заличаване на файловете му (освен тези в „Deleted Files“). Така не се налага прехвърляне на собственост на файлове, което може да бъде направено от конзолата или уеб интерфейса на облака. В допълнение, основните потребители, съхраняващи файловете, допринасят и за правилно

управление на дисковото пространство, като е възможно на потребителите, използвани от служителите, да се конфигурират квоти с цел ограничаване използването на частния облак не по предназначение.

При публични облаци необходимото пространство, с което е препоръчително потребител да разполага, може да бъде пресметнато по формула:

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

където: US – пространство на потребителя; UF – ресурси, които са собственост на потребителя.

За частни облаци, при които няма добавен модул за криптирано съхранение на файлове, формулата за изчисление на квотата на даден потребител би имала следния вид:

$$US = 2.5*FA,$$

където: US – пространство на потребителя (квота); FA – ресурси, до които потребителят има достъп.

За частни облаци, които са конфигурирани да съхраняват данните в криптиран вид на сървъра, формулата би била:

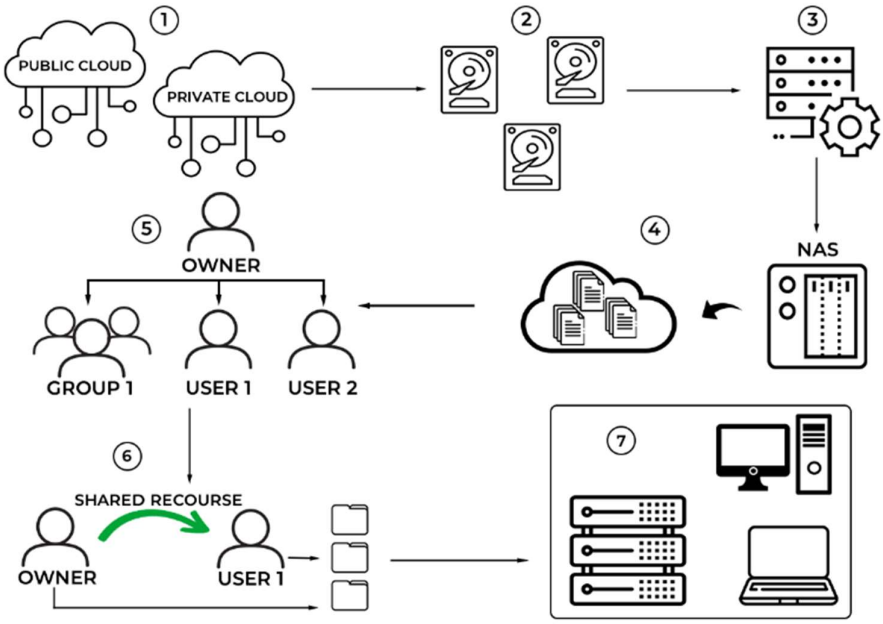
$$US= Ea*2.5*FA,$$

където: US – пространство на потребителя (квота); коефициент на увеличен размер на файловете вследствие на активиране на модул за криптирано съхранение на данните, където Ea > 1.0; FA – ресурси, до които потребителят има достъп.

Синтезирано, може да се заключи, че опорните точки за внедряване на частен облак в съществуваща ИТ инфраструктура по метода MICI (Method for Implementing Cloud in existing Infrastructure) са:

1. избор на вид облак;
2. изчисление на необходимите параметри за облачната инфраструктура – дисковете могат да бъдат изчислени по формулите: DS = D*(1+Cr) или DS= D*(Ea+Ea*Cr);
3. конфигуриране на избрания облак;
4. прехвърляне на съществуващите данни към облака;
5. добавяне на потребители и създаване на йерархия – квотите на потребителите могат да бъдат изчислени по формулите: US = 1,5*UF, US = 2.5*FA или US= Ea*2.5*FA;
6. споделяне на информацията с новите потребители;
7. конфигуриране на достъп на работните станции до облака.

Авторският метод MICI е подробно описан в подточка 2.4 от дисертационния труд.



Фигура 3 Метод за внедряване на облак в съществуваща инфраструктура

Авторският метод MICI е приложен в повече от 10 реални ИТ инфраструктури, които са различни по големина и сфери на дейност [Hristev, 2021]. Най-голямата организация, в която е интегриран частен облак по метода и е преминато от работа с NAS към NextCloud, е с над 300 работни станции. Методът е имплементиран и за миграция на малки фирми с под 10 работни станции, използващи споделени директории към частен облак. Част от организациите, при които данните са успешно прехвърлени в частен облак, са: Факултет по математика и информатика към ПУ „Паисий Хилендарски“, УМБАЛ „Каспела“, Мебелна фабрика „Колорадо“, Печатница „РайтПак“ ЕООД и други.

При направената литературна справка не са открити други подобни методи, описващи стъпките за преминаване от стандартна инфраструктура за съхранение на данни към облачно съхранение на информацията, съхранявана в ИТ инфраструктурите.

Резултати и изводи

Преминаването в дистанционен режим на работа, вследствие на Пандемията от Covid-19, накара множество корпорации в световен мащаб много бързо да реорганизиран установения работен режим. В същото време разработчиците на зловреден софтуер стават все по-изобретателни, за да могат да реализират своите цели.

От **Глава 2** на дисертационното изследване могат да бъдат синтезирани следните изводи:

- При стандартните инфраструктури за обработка на информационни масиви данните се разпространяват основно в локална компютърна мрежа.
- С обявяването на Пандемията от Covid-19, в началото на 2020 година, многократно са се увеличили справките за zero day exploits, на основните услуги, които позволяват отдалечен достъп до корпоративните частни мрежи.

- Докато стандартните методи за съхранение и обработка на информация използват файлова система за запазване на данните, облачните инфраструктури използват комбинация от файлови системи и база от данни.
- По-старите операционни системи и приложни програми е възможно да не поддържат новите мрежови протоколи, използвани от облаците за достъпване на съхранените на тях данни.
- Облачните инфраструктури поддържат различни и гъвкави начини за достъпване на данните.
- Преминаването от стандартна инфраструктура за съхранение на данните към облачна може да бъде сложна и комплексна задача.

В **Глава 2** са постигнати следните резултати:

1. Анализирани са методите за достъпване на данни при облачни инфраструктури.
2. Дефинирани и описани са подходите, позволяващи на по-стари операционни системи като Microsoft Windows 7 и Microsoft Windows 8 работа с данни, съхранявани в облачни инфраструктури.
3. Анализиран е хибриден модел на работа с облачни инфраструктури, който цели намаляване на мрежовата натовареност на инфраструктурата.
4. Създаден е авторски метод MICI за преминаване от стандартна инфраструктура към облачно съхранение на данните. По разработения метод са интегрирани облаци в повече от 10 ИТ инфраструктури, различни по обем и сфера на дейност.
5. Дефинирани са начините за изчисление на минимално необходимите налични ресурси в облачната инфраструктура, позволяващи възстановяване на информационни масиви след загуба.

Глава III. Възстановяване на изтрити и презаписани данни в ИТ инфраструктура

В Глава 3 от дисертационния труд са разгледани възможностите за възстановяване на данни, които са били изтрити или е записана нова версия на даден файл.

Възстановяване на данни при стандартни инфраструктури

В инфраструктурите, които използват стандартни методи за съхранение на данни като споделени директории, NAS и SAN, данните обикновено се съхраняват в различни видове файлови системи, определени от спецификата на инфраструктурата. При работата на твърдите дискови устройства, файловите системи съхраняват данните във фрагментиран вид. Тъй като не е ясно предварително какъв ще бъде размерът на даден файл, а и е напълно възможно той да увеличи размера си, то данните се записват на първия свободен сектор във физическия диск. Това означава, че един файл е разделен на парчета, които са записани на различни места в плочата на диска. Фрагментацията на данните води до значително забавяне при операции като четене и модификация на съществуващи файлове. Поради тази причина, когато системата не е натоварена, по голямата част от файловите системи автоматично дефрагментират данните или всъщност подреждат данните последователно, което би ускорило процесът по четене на файлове.

Когато даден ресурс бъде изтрит, докато не премине процес по дефрагментация, освободените сектори в твърдия диск биват маркирани като свободни. От гледна точка на файловата система това означава, че са сектори, на които може да бъде записана нова информация. При стандартната работа на компютъра постоянно се генерират файлове като различни кешове и други, които биха се записали веднага на свободните сектори.

При стандартните методи за съхранение на данни е възможно да се направи опит за възстановяване на информация чрез използването на софтуери за дълбоко сканиране на твърдите дискове. Те правят проверка на диска сектор по сектор, като се опитват да възстановят информацията за даден ресурс, ако са открити метаданни за него. Достатъчно е да бъде презаписана информация върху един от секторите на твърдия диск, върху който е имало информация на даден ресурс, за да бъде нарушена целостта на съответния ресурс. Обикновено това е възможно да се случи дори само да работи компютърът.

Дори в инфраструктурата да са имплементирани бекъпи по метода 3-2-1, който към момента е един от най-сигурните, то пак ще бъдат загубени данните от последния бекъп до момента на инфектирането на системата. Друг голям недостатък на този метод за съхранение на данни, който може да се отбележи, е, че по голямата част от фирмите не разполагат с разписани сценарии за действие при загуба на данни, а там, където ги има – рядко се отделя ресурс за проверка на целостта на архива и възможностите за възстановяване на системата до работещо ниво от архива.

Възстановяване на изтрити данни при облачни инфраструктури

Облаците разполагат с вградени функционалности като двойно изтриване на файловете, които могат да бъдат използвани за възстановяване на изтрита информация. Вземайки предвид спецификата на работа на облаците, дисковото пространство, което могат да заемат изтритите файлове, не е повече от 50% от квотата на дадения потребител. Периодът по подразбиране при частни облаци е 30 дни. Правилото е първият влязъл в изтритите файлове да бъде първият изтрит. Необходимото дисково пространство и потребителските квоти могат да бъдат изчислени по формулите от подпараграф 2.4 на дисертационния труд.

RDSC – авторски метод за възстановяване на изтрити данни, съхранявани в облачна среда - може да се заключи, че авторския метод RDSC за възстановяване на информационни масиви след изтриване на данни е следният [Hristev, 2022a]:

1. Идентифициране на компрометираното звено в инфраструктурата
   - В случай на Ransomware атака, първо трябва да бъдат открити колко работни станции в инфраструктурата са компрометирани и заразени с crypt ransomware. В зависимост от атаката това обикновено е един или група от компютри от определен отдел. Не е изключено да съществуват заразени компютри в повече такива.
2. Изолиране на компрометираните звена
   - При недобросъвестен служител трябва да бъде ограничен достъпът на служителя до системата.

• В случай на атака със зловреден софтуер, след като бъдат идентифицирани инфектираните компютри, те трябва да бъдат изолирани. Пропускането на тази стъпка би довело до цикличност в метода, защото след възстановяване на файловете чрез частния облак компрометираните устройства ще криптират отново възстановените файлове.

3. Изчистване на компрометираните устройства

   • В случай на зловреден софтуер компрометираните устройства трябва да бъдат изчистени от криптовируса. Компаниите, които разработват антивирусен софтуер, инвестират все повече средства за обновления, които да изчистят компютърните системи от подобни заплахи. След като бъде идентифициран вирусът, то може да бъде направена справка в интернет относно това кои продукти могат да засекат заразата и да я изчистят. Ако не бъде открит инструмент, чрез който да бъде отстранен проблемът, винаги съществува вариант заразените машини да бъдат преинсталирани.

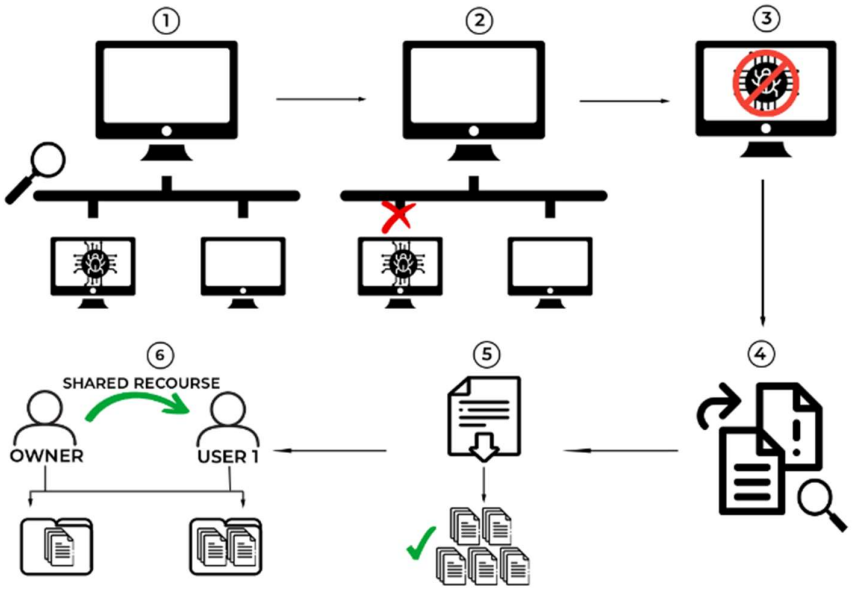
4. Възстановяване на изтритите данни от “Deleted Files”

   • Ако изтритите данни са били лични за потребителя, то те ще запазят структурата си и ще бъдат възстановени на местата, от които са изтрити.

   • Ако изтритите файлове са били споделени за потребителя, то те ще бъдат възстановени като лични за него, запазвайки първоначалната си структура.

5. Проверка на структурата на облака: структурата трябва да бъде проверена и ако файловете са изтрити в резултат на проникване на криптовирус, да се провери дали на първоначалното място на файловете няма качени криптирани файлове от криптовируса. Ако съществуват такива файлове, то те трябва да бъдат изтрити веднъж от хранилището, което е достъпно до потребителите, и втори път от “Deleted Files”.

6. Възстановяване на файловете на първоначалните локации и конфигуриране на достъпите до тях, ако се налага.



Фигура 4 Авторски метод за възстановяване на изтрити файлове съхранявани в облачна среда

Благодарение на авторския метод за възстановяване на изтрити информационни масиви RDSC (Recovery Data Stored in Cloud) бяха възстановени данни след заразяване с криптовируси в контролирани среди, а именно:

1. потребителски данни от контролирана Windows работна станция след заразяване с CERBER [Hristev, 2022b];
2. потребителски данни от контролирана Linux базирана Debian работна станция след заразяване с GonnaCry [Hrıstev, 2022a];
3. възстановяване на WordPress сайт и MySQL база от данни, работещи на контролиран Linux Debian сървър след заразяване с GonnaCry [Golev, 2022];
4. възстановяване на .Net Core приложение и архиви на бази от данни на Microsoft SQL Server, работещи на контролиран Windows Server 2019 след заразяване със Sodinokibi [Hristev, 2023];
5. неколкократни възстановявания на различни данни, изтривани от недобросъвестни служители в различни компании.

Направената литературна справка не откри други описани методи за възстановяване на изтрити данни, съхранявани в облачна инфраструктура.

Скриптове за автоматизирано възстановяване на изтрити данни, съхранявани в частен облак по метода RDSC – възможно е да се наблюдава забавяне при работата на уеб базираната част на приложението [Hristev, 2023]. Това, от своя страна, би довело до затруднения при възстановяване на криптираните файлове и изтриване на криптираните файлове от криптовируса и би отнело много време за възстановяване. Частните облаци NextCloud и OwnCloud имат сложна структура, като всички ресурси, налични за потребителя, освен във файловата структура на сървъра, биват описани допълнително и в базата данни на сървъра [Hristev, 2023]. За по-лесно управление и двата облака освен уеб базиран интерфейс разполагат и с OCC [Nextcloud, OCC command, 2023], чрез която част от функциите са достъпни и от шел. Това може значително да улесни администрирането им. Въпреки многобройните заявки от потребители за разработване на функционалност за OCC, с която да бъдат възстановявани файлове от Deleted Files, все още няма разработена такава [Hristev, 2023]. Създаден е скрипт, който позволява автоматизирано възстановяване на файлове по метода RDSC. Скриптът, описан в подточка 3.2.6 от дисертационния труд, след леки модификации, които са в зависимост от вида облак, ще позволи възстановяване на предходни версии на файлове, съхранявани в публични облаци. За да се реализира възстановяването е необходим административен достъп до сървърите, на които е инсталиран и работи публичният облак.

Подобрени скриптове за автоматизирано възстановяване на изтрити данни съхранявани в частен облак по метода RDSC - Скриптът, описан в подпараграф 3.2.6, ще възстанови всички файлове, които се намират в “Deleted Files” на потребителя. Това в част от случаите би довело до възстановяване на излишна информация, която потребителите в последствие трябва да изтрият ръчно. Именно поради тази причина скриптът е ревизиран и е предложена нова подобрена версия за

автоматизирано възстановяване на изтрити данни. В нея той приема като аргумент дати, между които ще се търсят файлове за възстановяване.

Възстановяване на презаписани данни

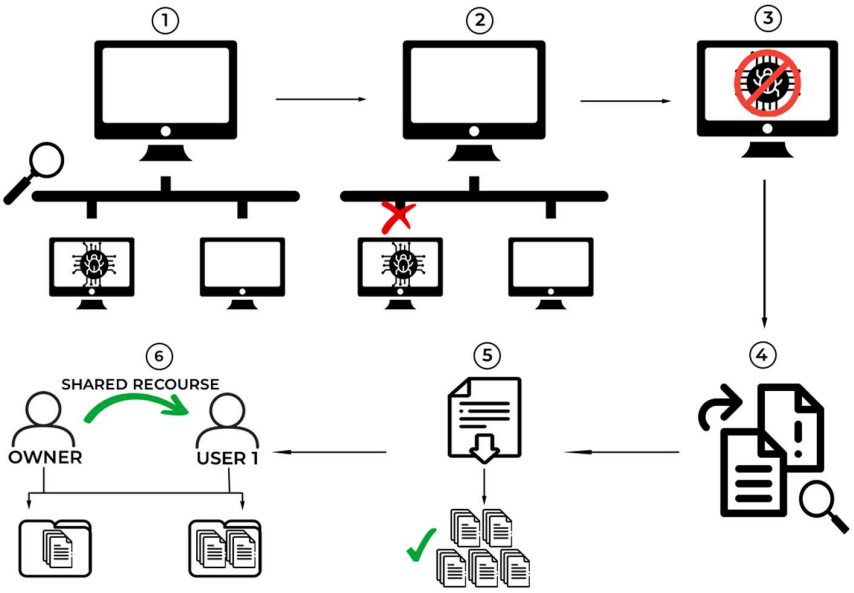
Разгледаните частни облаци и част от доставчиците на публични облачни услуги имат вградена функционалност за контрол на версиите на файл, който може да се използва при файл с нарушена цялост или е бил презаписан от недобросъвестни ползватели на системата. В допълнение, съществуват и криптовируси, които не изтриват файловете на системата, а всъщност криптират само част от метаданните на файла без да изтриват първоизточника. В такива ситуации информационните масиви могат да се възстановят посредством контрола на версиите.

RVSC – авторски метод за възстановяване на презаписани данни, съхранявани облачна среда - облаците пазят различни версии за файловете, съхранявани на тях. По подразбиране нямат зададени параметри относно това колко време назад да се съхраняват версии за съответните файлове. Ограничението във версиите се основава на базата на квотите, като размерът на съхраняваните версии не може да бъде повече от 50% от свободното пространство от квотата на потребителя [Nextcloud, File Versions, 2023].

Авторският метод RVSC (Recovery Versions Stored in Cloud) за възстановяване на презаписани данни, съхранявани в облачна среда, е следният:

1. Идентифициране на компрометираното звено в инфраструктурата.
   - В случай на грешка от потребител е необходимо да се идентифицира дали презаписването на данни е било умишлено или не е. Ако презаписването е по погрешка, то може да се премине към стъпка 4.
   - В случай, че презаписването е вследствие на ransomware атака, то първо трябва да се открият колко работни станции в инфраструктурата са компрометирани и заразени с crypt ransomware. В зависимост от атаката това обикновено е един или група от компютри от определен отдел. Не е изключено да съществуват заразени компютри в повече от един отдел.
2. Изолиране на компрометираните звена.
   - При недобросъвестен служител първо трябва да се ограничи достъпа на служителя до системата.
   - В случай на атака със зловреден софтуер, след като бъдат идентифицирани инфектираните компютри, те трябва да се изолират. Пропускането на тази стъпка би довело до цикличност в метода, защото след възстановяване на файловете чрез частния облак компрометираните устройства ще криптират отново възстановените версии.
3. Изчистване на компрометираните устройства – при зловреден софтуер компрометираните устройства трябва да бъдат изчистени от криптовируса.
4. Идентифициране на презаписаните файлове.
5. Изтегляне на предходни версии на презаписаните файлове и проверка на целостта на данните в свалените версии.

6. Възстановяване на презаписаните файлове в първоначалните локации и възстановяване на достъпите до тях, ако файловете са били споделени.



Фигура 5 Авторски метод за възстановяване на презаписани версии на файлове, синхронизирани с облачна инфраструктура

Благодарение на авторския метод RVSC, многократно са възстановявани потребителски данни, презаписани вследствие на небрежност или незнание на служители в различни по вид и обем организации. Методът за възстановяване може да бъде използван и за възстановяване на данни, когато има прекъсване на електрозахранване и загуба на информация в дадени файлове.

Направената литературна справка не откри други описани методи за възстановяване на презаписани данни, съхранявани в облачна инфраструктура.

Скрипт за автоматизирано възстановяване на предходни версии на файлове, съхранявани в частен облак по метода RVSC - възстановяването на презаписани данни може да се окаже много трудоемка задача, тъй като възстановяването на версиите се случва файл по файл и няма реализирани функционалности за подобно действие върху множество файлове едновременно. Освен това е препоръчително файловете да не се възстановяват директно, а първо да се свалят и да се проверят наличните данни във файла. Създаден е скрипт, който позволява автоматизирано възстановяване на файлове по метода RVSC. Скриптът приема като аргументи време и потребител, като ще възстанови най-близката версия на всички файлове за съответния потребител. За приключване на възстановяването на файлове по метода RDVC, е необходимо файловете да се преместят на първоначалните им локации и ако се наложи, да се конфигурират достъпите до тях. Скриптът, описан в подточка 3.3.2 от дисертационния труд, след леки модификации, които са в зависимост от вида облак, ще позволи възстановяване на предходни версии на файлове, съхранявани в публични облаци. За да се реализира възстановяването е необходим административен достъп до сървърите, на които е инсталиран и работи публичният облак.

Резултати и изводи

Финансовата стойност на данните, съхранявани в ИТ инфраструктура, надхвърля многократно стойността на хардуера, необходим за съхранението им. Данните, от своя страна, са сред уязвимите ресурси. Освен стандартните заплахи над тях допълнително се развиват и разработчиците на зловреден софтуер. През последните години атаките, насочени към информационните масиви, стават все по-големи като обем. След навлизането на RaaS на пазара в края на 2020 година се наблюдава повишаване на атаки с криптовируси с над 50% в световен мащаб.

От **Глава 3** на дисертационното изследване могат да бъдат синтезирани следните изводи:

- Софтуерите за дълбоко сканиране на твърдите дискови устройства не винаги успяват да възстановят загубена информация, заради спецификата на работата на файловите системи.
- Създаването на резервни копия на информацията не винаги е достатъчно гъвкаво, като те също могат да бъдат изтрити или компрометирани.
- При по-голямата част от корпорациите няма разписани сценарии за действие при загуба на данни, а там, където ги има – рядко се отделя ресурс за проверка на целостта на архива и възможностите за възстановяване на системата до работещо ниво от архива.
- При стандартните инфраструктури за съхранение на данни е почти невъзможно възстановяването на файлове с променена в тях информация.
- Облачните инфраструктури разполагат с вградени функционалности, които могат да подпомогнат възстановяването на изтрита или променена информация при масивите от данни.
- Частните облаци могат лесно да бъдат конфигурирани да покриват стандарт ISO 27001.

В **Глава 3** са постигнати следните резултати:

1. Проучени са и са анализирани подходите за възстановяване на информационни масиви, съхранявани в стандартни и облачни инфраструктури.
2. Създаден е авторски метод RDSC за възстановяване на изтрити данни, съхранявани в облачна среда. По метода RDSC са успешно възстановени:
   - потребителски данни, съхранявани на Windows базирана работна станция, след заразяване с криптовирус CERBER;
   - потребителски данни, съхранявани на Linux Debian базирана работна станция, след заразяване с криптовирус GonnaCry;
   - приложен сървър и СУБД, работещи на Linux Debian сървър, след заразяване с криптовирус GonnaCry;
   - приложен сървър и СУБД, работещи на Windows Server 2019, след заразяване с криптовирус Sodinokibi;
   - неколкократни възстановявания на различни данни, изтривани от недобросъвестни служители в различни по вид и обем компании.

3. Създадени са авторски скриптове за автоматизирано възстановяване на изтрити информационни масиви по метода RDSC. Авторските скриптове са подобрени с добавянето на времеви интервал за възстановяване.
4. Създаден е авторски метод RVSC за възстановяване на предходни версии на файлове, съхранявани в облачна среда. По метода RVSC многократно са възстановявани потребителски данни, които са презаписани вследствие на небрежност или незнание на служители в различни по вид и обем организации.
5. Създаден е авторски скрипт за възстановяване на предходни версии на файлове по метода RVSC с избор на времеви диапазон.

Заключение

Вследствие от направеното проучване в дисертационния труд са постигнати следните резултати:

- Направен е обзор на методите и техниките за съхранение и обработка на информационни масиви в съвременните ИТ инфраструктури.
- Изследвани са видовете заплахи за данните, съхранявани в ИТ инфраструктури.
- Анализирани са възможностите за възстановяване на изтрити данни и предходни версии на файлове.

Създаден е авторски метод MICI за преминаване от стандартна инфраструктура към облачно съхранение на данните. По метода са интегрирани облаци в повече от 10 ИТ инфраструктури. Част от организациите, при които данните са успешно прехвърлени в частен облак, са: Факултет по математика и информатика към ПУ „Паисий Хилендарски“, УМБАЛ „Каспела“, мебелна фабрика „Колорадо“, печатница „РайтПак“ ЕООД и други. Анализиран е хибриден модел на работа с облачни инфраструктури, който цели намаляване на мрежовата натовареност на инфраструктурата.

Анализирани са обичайните заплахи, водещи до загуба на данни в ИТ инфраструктури. Изследвани са стандартните методи за разпространение на зловреден софтуер и са класифицирани три вида изнудващ зловреден софтуер.

Създаден е авторски метод RDSC, на база на метода е написан скрипт, който автоматизирано възстановява изтритите и презаписани данни при частни облаци. Той е подобрен чрез добавяне на времеви критерии за данните. По метода RDSC успешно са възстановени:

- потребителски данни от контролирана Windows работна станция след заразяване с CERBER;
- потребителски данни от контролирана Linux базирана Debian работна станция след заразяване с GonnaCry;
- възстановяване на WordPress сайт и MySQL база от данни, работещи на контролиран Linux Debian сървър, след заразяване с GonnaCry;
- възстановяване на .Net Core приложение и архиви на бази от данни на Microsoft SQL Server, работещи на контролиран Windows Server 2019, след заразяване със Sodinokibi,

- неколкократни възстановявания на разнообразни данни, изтривани от недобросъвестни служители в различни компании.

Създаден е авторски метод RVSC, на база на метода са реализирани авторски скриптове, които автоматизирано възстановяват предходни версии на данни, съхранявани в частни облаци, чрез задаването на времеви критерии. Позовавайки се на метода, многократно са възстановявани потребителски данни, които са били презаписани поради небрежност или незнание на служители в различни по вид и обем организации, в които има интегриран частен облак.

Перспективи за развитие

С навлизането на частните облаци и нарастващите размери на информационните масиви, циркулиращи в ИТ инфраструктурите, неимоверно разработените авторски скриптове за автоматизирано възстановяване на данни ще намират все по голямо приложение. Логично продължение на положените основи в дисертационния труд би било:

1. Разширяване на автоматизираните скриптове за възстановяване по методите RDSC и RVSC с проверка на мета данните в СУБД на частните облаци, за да е възможно възстановяване на данните на първоначалните локации чрез запазване на правата за достъп и споделянията на ресурсите.
2. Създаване на модули за NextCloud и ownCloud, позволяващи възстановяването на множество данни по методите RDSC и RVSC през уеб базирания интерфейс на частните облаци.

Основни приноси на дисертационния труд

Основните приноси в дисертационния труд са:

I. Извършен е анализ на основните видовете заплахи, които водят до загуба на информационни масиви в съвременните ИТ инфраструктури.

II. Създаден е авторски метод MICI за преминаване от стандартна инфраструктура към облачно съхранение на данни.

III. Разработен е авторски метод RDSC за възстановяване на изтрити и презаписани данни, съхранявани в облачна инфраструктура. Създадени са скриптове за автоматизиране на метода.

IV. Разработен е авторски метод RVSC за възстановяване на предходни версии на файлове, съхранявани в облачна инфраструктура. Създадени са скриптове за автоматизиране на метода.

Връзките между приносите, целите, задачите, мястото на описание в дисертационния труд и направените публикации са следните:

Принос	Цел	Задачи	Глава	Публикации
I	Б	1.	1	3
II	А	2.	2	1
III	Б	3	3	2, 3
IV	Б	4	3	

Публикации по дисертационния труд

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

Апробация

Резултати, получени в изследването, са използвани в следните международни, национални и университетски проекти:

• Проект МУ21-ФМИ-007 „Симбиоза между математиката и информатиката (СМИ във ФМИ)“ към НПД на ПУ, университетски, 2021-2022;

• Проект МУ21-ФМИ-009 „Систематични иновационни изследвания в математиката и информатиката“ към НПД на ПУ, университетски, 2021-2022;

• Проект МУПД23-ФМИ-009 „Развитие на ИКТ чрез нови изследвания и технологични решения“ към НПД на ПУ, университетски, 2023-2024;

• Проект КП-06 НП62/1 „Математическо и информационно моделиране на динамични процеси – нови теоретични резултати, методи за изследвания и приложения“ към Фонд „Научни изследвания“ на БАН, национален, 2022-2025.

Част от резултатите, получени в дисертационния труд, са докладвани на следните международни и национални конференции и семинари:

• 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

Резултатите, получени в дисертационния труд, са частично финансирани по Проект № КП-06 ПН 62/1, „Математическо и информационно моделиране на динамични процеси - нови теоретични резултати, методи за изследвания и приложения“, 2022 -2025.

Библиография

[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 писменехъ", 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, последен достъп октомври 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, последен достъп октомври 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, последен достъп октомври 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, последен достъп октомври 2023