Operational Resilience

Backup Strategies for Crisis Management in the Age of Ransomware

Master Thesis

Dorukhan Yesilli

Delft University of Technology

Operational Resilience

Backup Strategies for Crisis Management in the Age of Ransomware

by

Dorukhan Yesilli

in partial fulfilment of the requirements for the degree of
Master of Science
in Engineering and Policy Analysis
at the Delft University of Technology,
to be defended publicly on August 31, 2023 at 12:00 PM.

Student number:	5539501	
Project duration:	March 01, 2023 – August 31, 2023	
Thesis committee:	Prof. Dr. M. van Eeten,	TU Delft, Chair, Section: O&G
	Dr. Y. Zhauniarovich,	TU Delft, First Supervisor, Section: O&G
	Dr. J. Ubacht,	TU Delft, Second Supervisor, Section: ICT
	LLM. L. Willinge,	EY, External Advisor
	ir. R. Boomsma,	EY, External Advisor

Cover: A very large array of electronic equipment in a room by
Matthieu Beaumont, Unsplash

An electronic version of this thesis is available at
https://repository.tudelft.nl/islandora/object/uuid%3A17e68d16-0918-4592-8f27-aa36b864fc98?collection=education.

Executive Summary

Cyber threats are becoming increasingly common, and organizations are being consistently targeted. One of these cyber threats, which represented 17% of all cyberattacks in 2022 worldwide, is Ransomware [1]. Ransomware is malicious software that infiltrates systems and is characterized by it encrypting files and leaving a ransom note behind, noting that the organization can get their files back if they pay. However, even paying in this circumstance does not guarantee that the victim will get their data back. These attackers do double and triple extortion methods where they exfiltrate the data and blackmail the victim again, noting that they will leak or sell organizational data, or they blackmail the investors/customers of the victim to make them comply and pay to the attackers.

Ransomware operate most notably in five main stages. These stages are the initial access, reconnaissance, lateral movement and privilege escalation, and impact. The behavior of the ransomware is quite undiscernible from other types of malware until they start encryption processes. They utilize encryption once they have enough access to impact the organization significantly.

Main ransomware targets are critical infrastructure systems, such as healthcare, governmental entities, financial entities, oil companies, electricity companies, mail delivery systems, etc., where the victims would be more willing to pay. The FBI notes that the potential loss from ransomware incidents is about $10.2 billion per annum in 2022 [2]. Not only that, but disruptions in critical infrastructure can lead to catastrophes that involve human lives and livelihoods.

A robust recovery mechanism for organizations is essential to mitigate disruptions to operations and downtime. Recovery could only be done if there are available backups and fit-for-use. The thesis precisely focuses on this issue. The thesis aims to establish insights regarding the best backup & recovery strategies in the field, what ransomware can do to circumvent them, and what challenges organizations face regarding backups & recovery. The subquestions directly relate to one of these three major points, and the main research question aims to explore what recommendations could be made to organizations in both private and public sectors, following the insights into the significant points in backup & recovery.

The research has been conducted on four types of data sources, creating a mix of primary and secondary data for a comprehensive overview. These sources include academic literature, cybersecurity frameworks, antivirus or backup solution provider reports, security blogs, and semi-structured interviews. The sources have been analyzed through two qualitative methods: semi-systematic literature review and qualitative content analysis.

The main findings of the ransomware attacks against backups shows that the name of the game, is access and, ransomware wins if it gets access to the backup files. The ransomware do reconnaissance through the organization's network to find network file shares that can connect to backup systems and access to backup files. Moreover, they do privilege escalation and lateral movement simultaneously by utilizing different tools and methods to further get inside the network, giving them increased privileges that allow them to corrupt, infect or delete backup files. They utilize vulnerabilities found in the backup solutions to increase their access level in the network further. Ransomware attacks also utilize legitimate service tools reserved for administrators, such as the Volume Shadow Copy Service, the Windows Management Instrumentation, etc., to tamper with the integrity of the backups. They also identify and terminate databases, antivirus, or backup software to release the backup data from processing so that they are not secure and can be encrypted. Lastly, if the ransomware cannot access the backups, they try to attack the management systems that create it. Here, they try to poison future backups, stay undetected in the system, and not start encryption unless they know the organization will face severe data loss. The poisoning method also works by infecting the backups so that the ransomware attackers can still have access to the system even if the system is rebuilt from backups.

The best backup & recovery practices that can protect against these ransomware attacks against backups are primarily categorized under; protection, generation, testing, and creation of correct backup & recovery policies. Organizations must protect the backup and recovery data through proper access and privilege management, monitoring, encrypting backups, establishing network controls, patch management, and disabling unnecessary services and protocols. On the generation side, organizations can try to create regular, redundant, air-gapped, immutable backups stored offsite and isolated offline. The organizations must also test these backups and their recovery practices to verify their availability and the strength of their incident response under a crisis scenario. The generation of backups, and their specifications including their protection, testing, and recovery processes, must all be explicitly defined and documented under backup & recovery policies.

The organizations, however, face many challenges when it comes to backup & recovery practices. These are mainly related to complexity management, inadequate security and backup policies, high costs and low allocation of resources to cybersecurity, insufficient testing and training, and the lack of incomplete or inaccessible backups. The organizations also run into issues with the solution mixes that they utilize. It is also noted that many organizations are pretty immature about their cybersecurity practices, but many are overconfident in their untested backups. They also carry a willingness to pay due to focus on core operations and ransomware insurances instead of allocating more resources for in-house or hybrid cybersecurity developments.

Considering the insights from the main three consideration points, which correspond to answering the subquestions of the thesis, a recommendation list is generated consisting of the findings given in the prior paragraphs. The recommendations for organizations are mapped out by comparing and contrasting the findings in ransomware attacks and best practices. Furthermore, recommendations are made on the challenges that organizations face. The recommendations will hopefully lead to organizations having better backup & recovery policies, plans, and processes established to ensure operational continuity.

Acknowledgements

At one point in my thesis, I started to think that I would not come to the point to write my acknowledgements. I have gone over much and had to face many negative situations during my time in these past 6 months external to my thesis. However, I can whole heartedly say, that I have managed with thanks to the people who always supported me and were there for me. I would like to list my thanks;

Firstly, I would like to thank my parents. My mother Feray Yesilli and my father Hakan Yesilli who have always put me first and ensured that I would have any support I need. They were there listening to my cries, my joys and were always there in my failures and successes. I greatly appreciate that both of you are in my life. Thank you.

Secondly, I give my thanks to my great committee members; Michel van Eeten, Yury Zhauniarovich, Jolien Ubacht, Ruurd Boomsma and Lukas Willinge. All of you have spent valuable time to give me feedback to encourage my learning and for me to create this valuable research. Although I will thank each of you seperately, know that I deeply value all your efforts in helping me. I also would like to thank Joy van Eijk who was always happy to help me to schedule appointments with everybody. Here, I also would like to thank Nicolas Dintzner as he has made time to ensure that my HREC application was done correctly.

Thirdly I would like to thank my friends. Life is only meaningful with our bonds to one another. I would like thank; Kursad Akdogan, Ludovica Bindi, Marion Dugue, Aspasia Panagioditou, Yaren Aslan, Helena Heijmen, Pascal Kampert, Alexandre Curley, Floris van der Heijde, Auriane Tecourt, Nicolo Canal, Maanav Jhatakia, Rhys Evans, Madalin Simion, Abdulkhamid Mukhamedov, Canan Tezgec, Nachiket Kondhalkar, Folkert Post, Gerdus van der Laarse, and Matvei Isaenko.

Lastly, I would like to thank the research participants who I cannot name due to ethical concerns, but a great part of my thesis was only possible with their contribution. I would like to specially thank Haiko van der Voort, Trivik Verma, and Ofke Teekens for helping me grow both as a researcher and as a professional. I would also like to thank other faculty staff who were always happy to provide support.

If I have unintentionally omitted your name, please accept my apologies, and know that I really appreciate your part in my studies and life.

Completing the thesis have allowed me to learn new skills especially regarding qualitative studies, of which I hadn't done extensively before. I have learned much regarding ransomware, backups, recovery and the challenges of organizations during this thesis. I have also learned much about myself. I found out that I had many skills and strengths that I did not know before. I also found out my limitations, and what I need to focus more on. I could only do this with the help of all the great people that have surrounded me.

Dorukhan Yesilli
Delft, August 2023

Contents

1 Introduction

1.1. Background

In our increasingly digitized world, cyber threats have become a pressing concern for individuals, businesses, and governments. It has become such a pressing concern that Ransomware is a grand challenge that is global in scale, extremely complex, concerning many stakeholders, and is ever-changing.

Ransomware is a growing but dangerous trend among the plethora of cyber threats. According to IBM Security X-Force Threat Intelligence Index 2023, ransomware attacks represented 17% of all cyberattacks worldwide in 2022 [1]. To understand the problem and the problem's relevance, it is essential first to understand what exactly Ransomware is and its impacts.

1.1.1. What is Ransomware?

Ransomware is malicious software (malware) designed to infiltrate systems and find and encrypt victims' files, including their backups. The attack then leads to the extortion stage, where the encrypted files are held hostage until a ransom payment has been made to the attackers [3] [4]. Even if the payment is made and encrypted files are recovered, the attackers may follow into double extortion, where they blackmail the victim after exfiltrating important classified data. Data exfiltration here means that the attackers take sensitive data outside of the organization and mean to either sell it or release it publicly. Even if the double extortion works and the victim pays the money, ransomware attackers may move to triple extortion, where they go to the clients of the victim organization and extort money from them as well. There of course, is no guarantee that the attacker will comply with their promises.

1.1.2. Ransomware Taxonomy

Ransomware, although consistent most of the time, is categorized differently in the literature. It is, however, important to note what characteristics are prominent and are recorded by researchers to understand and have insight into jargon used within the field. [5], [6] and [7] agree and categorize ransomware into three through their behavior. They note that some ransomware blocks access to the victim's systems and is named locker ransomware. If the ransomware uses encryption methods to hold the victim's files hostage, they note this as crypto-ransomware, and if both behaviors are prevalent, [5] notes this as an aggregate locker/crypto-ransomware. On the other hand, [8] and [9] go deeper and categorize ransomware based on severity, platform, and targets. [8] is different in adding a crypto-based category; and delineates ransomware regarding whether they do symmetric, asymmetric, or hybrid encryption. Both agree that; severity-based classifications are scareware, locker, and crypto-ransomware. Scareware here is a type of ransomware that scares the victim to buy or download unnecessary software. Regarding targets, [8] and [9] note consumer and organizational ransomware. Platform-wise, they note PC (Windows, Linux, etc.), Mobile, Cloud, IoT, and Ransomwear (targeting wearables like smartwatches) are the platforms targeted.

[10] gives the most comprehensive taxonomy that includes all the previous categorizations but also categorizes ransomware with respect to its infection, communication channels (establishment of command and control), and malicious actions. Under infection, they note phishing, malicious agent actions, drive-by-downloads, and vulnerabilities in systems. In context, phishing is a method where victims are contacted through phone, text, or email to either lure victims to provide sensitive information or to make them download malware. Malicious agents in this context relate to, e.g., disgruntled employees who may infect the systems within a company. Drive-by-downloads are about an employee clicking a non-secure site and accidentally downloading malware. Vulnerabilities are how malware exploits operating systems, browsers, or unpatched software to deliver their payloads. Payload in this context means what ransomware does in the victim's computer (and/or network). Command and control communication, on the other hand, is essential as ransomware can establish contact with the attackers through hard-coded IPs or Dynamic Domains where the latter is harder to detect. [11] notes that not all ransomware establish a command and control channel and are more independent. Lastly, [10] delves into malicious actions. Much like the previous mentions, it notes encryption and locking as two actions. However, it also notes data exfiltration as a third main action.

Thus, there is a lot of research into the taxonomy of ransomware, and researchers tend to categorize ransomware based on their research goals. In this thesis, the search findings will be skewed towards crypto-ransomware, among them the ones that primarily target organizations and take the actions of encryption, locking, and data exfiltration.

1.1.3. Ransomware Lifecycle

The ransomware lifecycle, just as taxonomy, is denoted differently in each research. Although the names change, there is consensus on the main behavior that it exhibits. According to [5], the five phases of Ransomware are: ”[vulnerability] exploitation & infection, [payload] delivery & execution, [file & backup] spoliation, file encryption, and person notification.”

According to [9], on the other hand, ransomware has 5 phases. The first stage is called the ”distribution phase,” and the ransomware is delivered into the victim's system through exploiting vulnerabilities or other infection methods. The second stage is the ”Reconnaissance Phase,” where the ransomware collects information on the victim's network. At this stage they will discover the operating system (OS) and the network specifications. The third stage is the ”Preparation Phase,” where ransomware searches for important files, resources, and privileged access to the network. It is also in this stage that, if applicable, ransomware contacts the command and control channel. Here it is noted as ”if applicable” because [12] states ransomware families such as SamSam do not communicate with a command and control server but do the encryption locally. Continuing with the stages from [9], the fourth stage is the ”Hijacking Phase,” where they start locking, deleting, altering, and encrypting important data and backups. Lastly, they move on to the extortion phase, where they ask for ransom.

For visualization, a good and simple anatomy of a ransomware attack is described in Figure 1.1. The diagram shows that the attacker first looks for ways into the network. Here, many tactics, techniques, and procedures (TTPs) could be utilized to find a way into the victim’s system. Once the attacker is in, they try to establish a link to their command and control channel. They then do lateral movement, which means they dig into the victim system to understand where important files and backups are located. They also try to escalate privileges to gain elevated access that a normal user would not have. They try tactics to do either vertical privilege escalation that would go from, for example, a user to an admin-level account, or horizontal privilege escalation where they have access to other data that another similar-level account has permissions to access. Lastly then, the data are exfiltrated, backups are destroyed and the data is encrypted.

In all three cases the main behavior of ransomware is established. In summary the operations of a ransomware goes like this: infect, learn about your victim’s system, escalate your privileges for access, find where the most critical files and backups are, exfiltrate & infect, encrypt and destroy files, then ask for ransom.

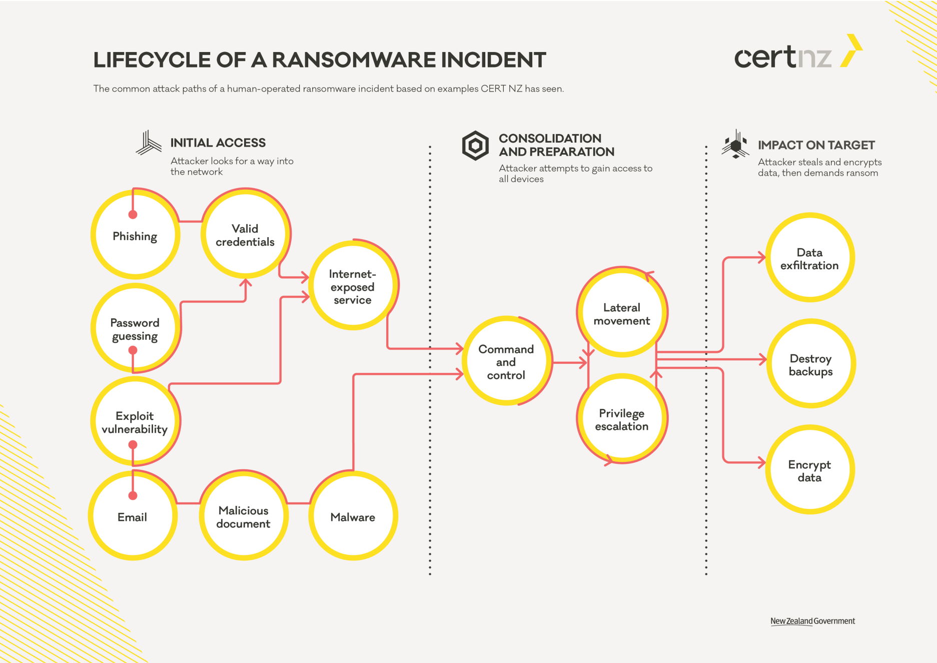
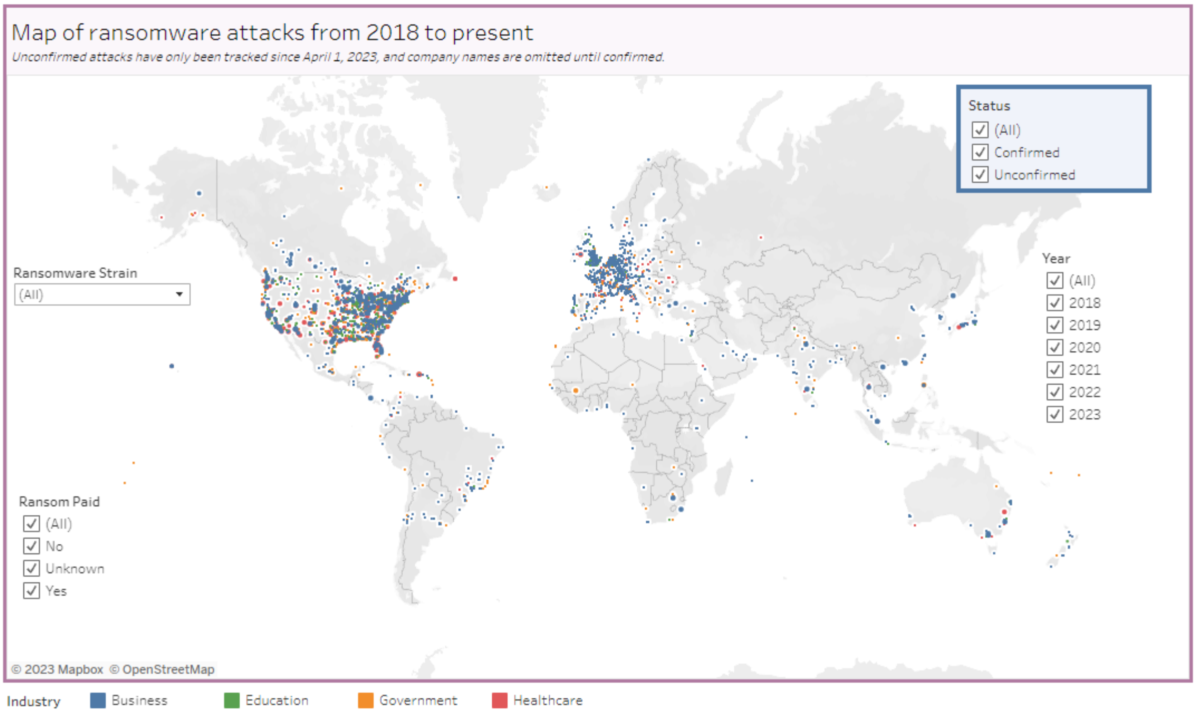

LIFECYCLE OF A RANSOMWARE INCIDENT
The common attack paths of a human-operated ransomware incident based on examples CERT NZ has seen.
certnz
INITIAL ACCESS
Attacker looks for a way into the network
CONSOLIDATION AND PREPARATION
Attacker attempts to gain access to all devices
IMPACT ON TARGET
Attacker steals and encrypts data, then demands ransom
Phishing
Valid credentials
Internet-exposed service
Password guessing
Exploit vulnerability
Email
Malicious document
Malware
Command and control
Lateral movement
Privilege escalation
Data exfiltration
Destroy backups
Encrypt data
New Zealand Government

Figure 1.1: Ransomware Anatomy from the New Zealand Cyber Security Authority [13]

1.1.4. Ransomware Targets

Ransomware primarily targets critical sectors where the necessity to pay the ransom would be higher. According to the US federal authorities (FBI) [2], Ransomware targeted the healthcare sector in 2022 more than any other critical infrastructure. [14] from Heimdal Security gives some examples of ransomware attacks where the attacks have been on the UK mail delivery system ”Royal Mail,” which is classified as a critical infrastructure, Tallahassee Memorial Healthcare in Floria, Hospital Clinic de Barcelona, and governmental entities. Aside from these examples, [15] keeps an updated and extensive list of the current ransomware attacks disclosed since 2018, categorizing the specific ransomware strains at play. Both lists contain examples to show how no organization on any level is safe and must be vigilant against Ransomware at all costs. Regarding the location, the ransomware threat is global; however, it is seen that ransomware threat is increasingly common in the US and Europe, according to Figure 1.2. Moreover, with blue dots consisting the majority, it could be identified that ransomware mostly target businesses.



Map of ransomware attacks from 2018 to present

Unconfirmed attacks have only been tracked since April 1, 2023, and company names are omitted until confirmed.

Status
☑ (All)
☑ Confirmed
☑ Unconfirmed

Ransomware Strain
(All) ▾

Year
☑ (All)
☑ 2018
☑ 2019
☑ 2020
☑ 2021
☑ 2022
☑ 2023

Ransom Paid
☑ (All)
☑ No
☑ Unknown
☑ Yes

Industry 🟦 Business 🟩 Education 🟧 Government 🟥 Healthcare

Figure 1.2: Ransomware Map from [15]

1.1.5. Ransomware Impacts

Ransomware attacks lead to significant economic loss for both organizations and consumers. The FBI notes that the potential loss from these ransomware incidents is about 10.2 billion $ in 2022 [2]. According to [16], by 2031, the damage costs are predicted to exceed $ 265 Billion without even considering the new estimates that have been published by Verizon [17]. For example, UK's "Royal Mail" was asked for a ransom equivalent to $75 million [15]. The impacts, however, are not limited to economic losses. There is also a high probability of temporary or permanent loss of sensitive information that must be kept secret to comply with respective regulations (e.g., GDPR). This can harm an organization's reputation, hindering its ability to retain customers and partners. The ransomware attacks also lead to unexpected and long downtimes. This is highly critical in sectors such as Healthcare, Energy, Delivery, etc., where the downtime in operations can lead to cascading effects that may not even be locally contained, and spill into other sectors. An example to highlight this spill was when the Colonial Pipeline Company had a downtime of 5 days to restart regular operations. Even though they paid the ransom to become operational again, the downtime caused widespread fuel shortages, panic-buying at the pumps, and a spike in gas prices, causing even airlines having to change their flight schedules [18]. Thus, considering the vast impacts Ransomware can cause, it is crucial to be aware and employ strategies to detect, prevent, and mitigate ransomware attacks.

1.1.6. Ransomware Mitigation

All is not lost however. There is much research and solutions into the detection, prevention and mitigation of Ransomware attacks. The problem is, that most of these solutions come with a price tag, and with their own challenges. [19] notes that the best protection in the literature is always done by a combination of detection and backup policies. Backups are the final card to be played against ransomware. Backups are important as the moment you receive the ransom note, it is already too late and your systems are already compromized. In that moment, knowing that the organization has a safe backup that had been updated frequently and tested prior will be the breaking point.

Backups entail copying physical or virtual files or databases to another location for preservation in case of crisis [20]. Data backup solutions can have physical, virtual, cloud, mobile, database, and copy data management properties that could be considered. According to some, these data backups become a prevention strategy and, according to others, an impact mitigation strategy against ransomware. If the critical data and infrastructure are managed well, ransomware attacks would lose their primary business model and only become a nuisance for organizations with proper backups. Thus, organizations require strong backup policies and solutions.

1.1.7. Problem Statement

Although the mitigation section provides some comfort with the existence of backups; Work-as-intended seldom matches how the real-world works. It is important to identify the interconnected nature of ransomware and backups, as a growing number of ransomware strains specifically target backups [21] [9]. This is done in order to make the victim more willing to pay, leaving them without a way to recover. Then it is essential to understand how ransomware circumvents protection on backups or destroys these backup files.

Moreover, in real life, there are many challenges organizations have with backups and recovery. This is due to reasons such as insufficient resources, low cybersecurity awareness, etc. [22] notes that most of the reasons for this insufficiency stem from organizations being focused on their core value offering and that they delegate a low amount of resources towards proper cybersecurity (and thus backups). For example, [23] notes that many healthcare sites did not have sufficient backups when they were struck by ransomware. [24] states that among the 374 attacks, they noted, only about 20% of the organizations were able to restore data from backups. Moreover, [25] notes that among the surveyed ransom victims, only 11% had recent backups.

Then, the problem could be summarized as: ransomware attacks backups, organizations need strong backup strategy and solutions, however organizations face challenges during recovery from backups.

1.2. Research Purpose

This research aims to establish insights regarding the best backup strategies in the field, what ransomware can do to circumvent them, and why recovery from backups is difficult in real-life cases. Then the insights will be the basis for creating a recommendation list regarding backup policies and establishing a holistic view for decision-makers thinking of strengthening their cybersecurity practices.

These insights will be generated from the reports of credible antivirus software, backup solutions providers, security blogs and established cybersecurity frameworks. Further discovery could also be made through industry insights, through expert interviews. Figure 1.3 summarizes how the mentioned research purpose is realized in the thesis.

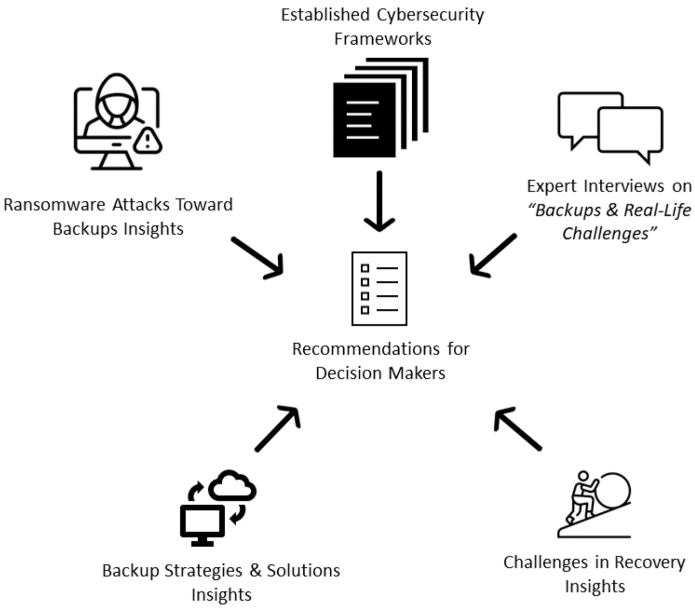

Established Cybersecurity Frameworks
Ransomware Attacks Toward Backups Insights
Expert Interviews on *"Backups & Real-Life Challenges"*
Recommendations for Decision Makers
Backup Strategies & Solutions Insights
Challenges in Recovery Insights

Figure 1.3: Thesis High-Level Overview

1.3. Research Gaps

The research gaps are identified under Section 2.3, which also serves as a conclusion to the semi-systematic literature review. The gaps identified could be summarized as;

- Ransomware attacks that are specifically scoped on backup and recovery is not well explored. There is a knowledge gap on how the current and most salient ransomware in the last 5 years, target backups and make recovery impossible. Most literature prior gives an overview of many aspects regarding ransomware and albeit identifying that there are attacks towards backups, they do not give extensive information on how the attacks are commenced with the specific goal of destroying backup integrity. There is also not an overview of what vulnerabilities of backup software or hardware was utilized to actually compromize backup data in the analyzed academical literature.

- There is a knowledge gap on cybersecurity frameworks applied in the context of backups & recovery in the academical literature. The question is; ”what do the cybersecurity frameworks recommend scoped on backups & recovery?” Although ransomware has been prior looked at through implementing the NIST Cybersecurity Framework and the CIS Controls; there is no insight into other relevant frameworks. For example there is no knowledge on how the new DORA regulations mandate EU based organizations to now be compliant with regards to backups & recovery, or how the Australian Cyber Security Centre (ACSC) has essential eight guidelines which also include actionable items for organizations in this context. Considering this knowledge gap, there is also no overview of all best practices regarding backup & recovery in the context of ransomware attacks.
- There is also a knowledge gap with regards to the difficulties and challenges that organizations face with the adoption and application of the best practices regarding backup & recovery. Albeit having examples of failure in recovery or unusuable backups, due to reasons such as not checking the validity, are mentioned in the literature; in the light of all the cybersecurity frameworks that are present in the industry, there is not enough insight into why organizations continue to fail recovery in the aftermath of a ransomware attack or why organizations have problems creating these backup & recovery policies and systems in the first place.

1.4. Research Objectives

Research objectives that follow the corresponding research gaps are:

- Explore currently known and most salient ransomware attacks on backups.
- Explore the state-of-the-art backup & recovery strategies utilized by organizations.
- Denote the challenges and difficulties that organizations face in backup & recovery.

The objectives are deemed achieved once there is enough data saturation, and that the findings start repeating themselves.

1.5. Research Scope

The thesis is aimed towards exploration however, no exploration will lead to outcomes without proper scope. The thesis positions itself as a high-level qualitative exploration of ransomware attacks, backup strategies and challenges of recovering from backups. For example, an attack on the management system of a tape-based backup can be described and how the attack is commenced by ransomware actors may be explained however, the vulnerabilities that have led to the attack will not be elaborated on other than their basic descriptions. This means that the thesis is not scoped to achieve deep technical information about these attacks.

Semi-systematic literature is scoped on research on ransomware, backups & recovery. Qualitative content analysis on the other hand focuses on documents (blogs, websites, written reports) published by credible antivirus software, and backup solution providers. Any other form of media are excluded. Documents will be purposefully selected from the web until data saturation is reached.

For backup strategies and challenges faced, the main samples will come from the cybersecurity frameworks that are established as best practices. Further analysis will be done with regards to reports published by credible antivirus companies & backup solution providers.

For the interviews, only people who have experience regarding at least one of the three focus objectives will be contacted. These people may be IT Experts, Forensics Experts, Red Teamers (people simulate attacks on organizations to find out weaknesses), Data Management and Protection Experts, Cybersecurity Consultants or etc.

1.6. Research Questions

The research questions are built upon the research gap, and the thesis aims to answer:

> ***What recommendations can be made for organizations to help them develop good backup and recovery strategies against the growing ransomware threat through a systematic and methodological approach?***

The research question is justified through the purpose of the thesis. The main deliverable of the thesis is a recommendation list that is generated through conducting a systematic and methodological research by utilizing the methods; semi-systematic literature review, qualitative content analysis, and semi-structured interviews. The research question follows the stated gap from [12], where the author had noted that a backup plan that denotes all information available is essential in the fight against ransomware. They had further explained that if there is lack in information, such gaps must be filled. There is a lack of information with regards to why organizations still have challenges with regards to the adoption of the best practices, and how ransomware currently target backups specifically. Thus, the research question aims to be a rigorous overview regarding the current information available on the topic. The subquestions which answer an aspect of the research question are:

- *SQ1: How do the most salient ransomware families target and compromise backups?*
- *SQ2: What are the current state-of-the-art backup & recovery practices that are (or can be) employed by companies to safeguard their data and ensure operational continuity?*
- *SQ3: What are the main challenges faced by organizations regarding backups & recovery?*
- *SQ4: What are the implications of the discovered findings with respect to providing a recommendation list to organizations?*

Here, it is seen that the subquestions manifest themselves through the previously identified objectives following the presented research gaps. Considering the previous argumentation on the research gaps the subquestions are justified in answering the research question as it is relevant knowledge for decision makers in organizations not only to know backup strategies but also to see; what are the best practices established by credible entities with regards to backup & recovery, what challenges organizations face during the adoption of these best practices, or what attacks ransomware utilize in order to make recovery impossible. Through this knowledge base the organizations will be able to make more informed and coherent choices with regards to their cybersecurity practices.

Considering the literature review, there is no work that brings and contrasts these focus areas to provide an overview on backup & recovery. The semi-systematic literature review, qualitative content analysis and expert interviews can highlight the trends and established practices in the domain. The sources utilized in order to answer the subquestions are sector-leading antivirus & backup solution provider reports, cybersecurity frameworks, security blogs, and expert interviews. The subquestions are thus the open ended exploratory questions that such a research needs. The last subquestion (SQ4) ensures that all the implications of the findings are translated into a recommendation list in Chapter 5. This research with respect to the main research question and its corresponding subquestions is gradually built upon, through the utilization of different methods which relate to the main chapters of the thesis.

1.7. Research Methodology

1.7.1. Research Design

The research design section discusses the philosophy, research type, research and sampling strategy, the time horizon, including the data collection and analysis methods.

Research Philosophy

It is imperative that the research philosophy, the research aims, objectives and questions match. The philosophy paves the way for the researcher and identifies the worldview of which they see through. The aim of the thesis is to establish insights on the best backup strategies, what ransomware do to circumvent them and why recovery from backups is so difficult in real life. This insight is then aimed to be utilized to create a recommendation list for organizations. The objectives are the manifestation of this aim and are thus open-ended exploratory objectives. The research question and sub questions operationalize the aims and objectives.

To this end, the research philosophy that has been utilized in this thesis is pragmatism. In a nutshell, pragmatism reveals a middle ground between positivism and interpretivism. It focuses on the usefulness and the applicability of the research findings rather than focusing into the philosophical position. [26] note that main sides that pragmatism reconciles here are positivism and constructivism. These methods favor quantitative (deductive) and qualitative (inductive) research designs respectively. Pragmatism on the other hand offers the researcher the ability to be able to select the research design and methodology that are most appropriate to address the research question rather than tangling in philosophical debates [26]. This is quite useful as the thesis utilizes qualitative content analysis and semi-structured interviews as two primary methods that will be triangulated in order to answer the research question and its corresponding sub questions. [27] notes that ”content analysis is a method that may be used with either qualitative or quantitative data and in an inductive or deductive way.” [28] provides insight on the adaptability of qualitative content analysis by noting that ”Qualitative methods allow the researcher to freely adapt the steps to the peculiar characters of the research project.” The same is true for semi-structured interviews, as it allows for open-ended conversation to tailor the questions to the interviewee. Considering such adaptability and a purposeful utility, the methods of the thesis match with the pragratism philosophy.

Pragmatism also fits the research question, the subquestions, the research aim and objectives. This is prevalent as the questions are basically the operationalization of the main deliverable, that is the recommendation list, for organizations. It is very utility based and is about collecting the established state-of-the-art knowledge. With this argumentation in place, it could be justified to say that this thesis is based on the pragmatic research worldview.

Research Type

The research is a qualitative and an inductive study. [29] notes that the qualitative research methods can utilize texts, in order to describe, summarize, examine or to develop a conceptualization of the content. This fits with the thesis as the subquestions aim to utilize text-based data from antivirus and backup solution providers, cybersecurity frameworks, and the statement or reports from ransomware victim organizations. These are all qualitative data and are to be summarized and conceptualized in order to give answers to each of the subquestions and the research question respectively.

The study is also inductive as the questions that the thesis answers are open-ended and exploratory. [30] argues that qualitative content analysis as opposed to quantitative content analysis utilizes open questions that guide research, and corresponds to inductive research. Qualitative content analysis makes the text an important driving source where the author will read through the textual data to identify concepts and patterns, including those that have not been foreshadowed through the literature review but emerge throughout the study. This way, the thesis is not limited to only the preestablished literature regarding ransomware attacks on backups for example. Inductive research will allow here to answer the subquestions whilst allowing the researcher to be flexible and open to new information. A deductive research on the other hand would have needed clear and narrow research questions that test certain hypotheses. It is more in line with using quantitative methods utilizing statistics etc. to be able to confirm theory or hypothesis.

Research Methods

There are three methods utilized within the thesis. One is semi-systematic literature review (Chapter 2), followed by a qualitative content analysis (Chapter 3), and lastly, semi-structured interviews (Chapter 4). Each method is explained in detail and fully in their own corresponding chapters. In a quick overview, each method tries to delve into answering the three subquestions and then the findings that are found over 3 different methods and data sources are then compiled and reflected upon in Chapter 5 in order to come up with policy recommendations for organizations which answers the fourth subquestion and consequently, the main research question.

The data collection has been done differently in each method. For systematic literature review the databses Scopus and ArXiv were utilized to conduct searches in order to find relevant academic articles with the main keywords ”ransomware,” ”backup,” ”data protection,” ”cyber recovery,” and ”challenge.” For qualitative content analysis the main data sources were established as: cybersecurity frameworks, reports from antivirus or backup software providers and their related cybersecurity blog articles a purposive sampling strategy has been utilized here in order to get information that is relevant to answer the questions of the thesis. For semi-structured interviews, again a purposive sample has been taken in order to find experts who have relevant experience in the domain of ransomware, backups & recovery.

The data analysis has been laid out in detail in each corresponding chapter of the methods. However in a general sense, semi-systematic literature review was screened two times prior to being fully read and analyzed in order to exclude irrelevant articles for ease of analysis. Following this, the remaining number of articles have been extensively studied, their quality has been noted with respect to whether they identify the topic correctly and whether they describe it in a objective and informing fashion. Following this, notes have been made to be used to create the findings of the literature review to find the relevant research gaps for the thesis and to answer the subquestions which have been iteratively generated following the identification of the gaps.

For qualitative content analysis the data analysis have been done in three stages. These are preperation, organization and reporting. The preperation phase selected the unit of analysis; where the unit can be words, sentences or paragraphs. After some trial runs a sentence-level unit was selected as it was suited to capture the meaning of a mentioned backup strategy, a challenge or a ransomware attack pattern. It is also not so granular that the idea flow is too fragmented, but it is not large to convelute the meaning of a backup strategy, etc. The analysis only decided to include manifest content, and therefore to identify information that is plainly given and is not implicit. Then open coding is done in the relevant material that has been collected, and the organization stage begins. During organization the qualitative data is organized under their respective categories and subcategories. The organized data is then made into a codebook and an abstraction is created which identifies categories and their corresponding subcategories. The reporting then is done by unloading content in a structured manner, also considering the trends and what the main consideration points in the documents were.

For semi-structured interviews the data analysis has been also done similar with qualitative content analysis. It followed the same preperation, organization and reporting structure in an identical fashion. Therefore all methods analyze and generate findings in order to answer the subquestions. In the end, all the findings are reflected upon by mapping the concepts together e.g. noting that one specific challenge may be circumvented by a best practice, etc. These reflections then create the main recommendation list which the main research question aims to deliver.

Sampling Strategy

The sampling strategy selected for both the documents (for content analysis) and interviewees are based upon non-probability sampling. Among the non-probability sampling strategies specifically the purposive sampling method is selected for both the semi-structured interviews and also the content analysis reports. Interviewees have been selected with regards to their expertise in the field and specifically with their experience with either backup strategies and solutions, recovery, or ransomware attacks. For the content analysis, the textual data has been sampled with regards to both discovered and advised knowledge about what avenues to search for. For example, many samples contain ransomware victim companies and their statements or other samples of data are taken specifically from the recommended cybersecurity frameworks. The sampling is also iterative as with new discoveries, new samples were generated to increase the rigor of the study. Further support is found in [30] as they argue that qualitative research may use ”purposive sampling to allow for identifying complete, accurate answers to research questions and presenting the big picture; selection of data may continue throughout the project”

Time Horizon

The time horizon selected for the research is cross-sectional. Therefore, every data point has been collected at one point in time. The reasoning behind this is the fact that neither the research aims nor the research questions aim to assess a change in perception or knowledge over time. However the research is answerable from data collected from a single point in time, which is also beneficial considering thesis time limitations, and interviewee availability.

1.7.2. Ethics

Regarding Ethics, qualitative content analysis does not face any issues as it is based upon published media.Semi-structured expert interviews on the other hand are a critical consideration.The procedure is designed with regards to strong ethical requirements that are based upon the mandatory TU Delft Research Ethics rules. Main considerations have been made with regards to the collection, storing and processing of personal and personally identifiable information data (PII). PII in this context means any sort of data that can be tied to the interviewee to discover who they are.

Storage of data has been secured with strong access management practices. The data has been stored in the TU Delft Institutional Storage and is kept private. There are four types of data that are generated through the semi-structured interviews. These are audio recordings, pseudonymized transcripts, technical summaries and signed informed consent forms. The technical summaries do not have any personal or PII data.

The only people with access to all data are the author (researcher) and the responsible researcher (first supervisor). The external advisors are only allowed access into the technical summaries that are generated through pseudonymized transcripts from the audio recordings that have been recorded during the interviews. The only pieces of data that are published are the (unsigned) informed consent form, and technical summaries which can be found under Appendix A and Appendix C respectively.

1.8. Engineering and Policy Analysis (EPA) & Academic Relevance

EPA is a Master's program that incorporates analytical thinking and exhibits the multi-actor perspective. The current topic is within a domain where different organizations and individuals have different views regarding the applications or the importance of cybersecurity in day-to-day business. This makes solution building and applications of these solutions quite complex. Complexity in such, according to [31], is best described by the understanding that this problem is not a problem with "...a single problem owner, a single problem definition, a small number of players and few alternative solutions" [31].

Cyber attacks do not only affect private organizations but also is relevant to any public organization on any level in any region connected to the network. As described in the introduction, critical infrastructures such as healthcare [23] and governmental entities are primary targets for these cyber attacks. It should be noted that this is the reason that the author utilizes always the word "organization" but not "enterprise" or "business" to emphasize generalizability.

Mitigation of the impacts of ransomware can be achieved by utilizing backup strategies. If organizations do not have much to lose, there is no hostage data and less necessity to pay. However, there is decision making on an organizational level with many different stakeholders who are part of other teams in different countries, who have various legislation and regulations to follow, it is not easy to implement strategies.

On the research approach, the qualitative content analysis and semi-structured interviews are relevant methods to the curriculum of EPA. Identifying key information from the knowledge base and triangulating with separate methods to achieve comprehensiveness and validity are skills that a student must be able to exhibit.

Therefore, the thesis is quite relevant to the EPA curriculum and is research that aims to explore the industry knowledge base to peck at this grand challenge. The backup strategies discovered from cybersecurity frameworks, and credible sources can then become a roadmap into how organizations can rethink their backup strategies and identify any gaps they have.

The Academic relevance of the paper comes from the fact that the thesis fills relevant research gaps to bring the research space a bit forward. It brings the industry expertise to explore and expand the understanding of a well established academical research topic. The understanding is expanded through comparing and contrasting attacks versus strategies, the notation of the real-life challenges, and identifying relevant cybersecurity frameworks to bring them into the awareness of researchers. The deliverable of the thesis could be building blocks for future researchers in the field as they can delve deeper to specific aspects presented in the thesis.

1.9. Structure of the Thesis

The structure of the thesis is as follows. First, the semi-structured literature review and their findings are given. This chapter also establishes the research gaps that have been found in the literature. The next chapter is qualitative content analysis. This is followed by the semi-structured interview findings. Each chapter that belongs to a different method share a similar template. Each show their findings about the ransomware attacks towards backups, notes the state-of-the-art backup & recovery strategies, and finishes with the challenges in backup & recovery. All method findings are then synthesized to generate one final recommendation list, generated from a discussion which compares the ransomware attacks, challenges in the domain and the backup strategies. Lastly, the conclusion, limitations and future research directions are presented.

2
Semi-Systematic Literature Review

Considering the research purpose, this chapter aims to explore the current state of academical research about ransomware attacks towards backups, backup & recovery strategies employed in the domain, and challenges faced during recovery by organizations by utilizing the semi-systematic literature review method. The chapter starts with the description of the method and the procedure that was followed, and then continues with literature findings on backups & recovery. Findings have been categorized under four main sections. This categorization was made after identifying the research gaps from the literature which serve as a basis for the research question and its corresponding sub questions. After the research question and the subquestions were determined, the findings have been collected and used as a primary input, that is aimed to be expanded through other methods, to produce policy recommendations for organizations in Chapter 5. Section 2.2.1 aims to answer the first subquestion, namely, *”How do most salient ransomware families target and compromise backups?”* Section 2.2.2 and Section 2.2.3 on the other hand aim to explore the second subquestion: *”What are the current state-of-the-art backup & recovery practices that are (or can be) employed by companies to safeguard their data and ensure operational continuity?”* Lastly, Section 2.2.4 explores the third subquestion: *”What are the main challenges faced by organizations regarding backups & recovery”* The expected output of each section is that it generates insights per subquestion which can be used as basis for reflection in Chapter 5 to produce policy recommendations which then leads to the answering of the main research question: *”What recommendations can be made for organizations to help them develop good backup and recovery strategies against the growing ransomware threat through a systematic and methodological approach?”*.

2.1. Description of the Method

The literature review has been conducted in a semi-systematic manner. The main guidelines of the PRISMA Protocol [32] were followed. The databases searched are Scopus and ArXiv. These were sufficient as Scopus also gives results from the ACM Digital Library and IEEE Xplore. Combined with the vast archive of ArXiv, enough literature to highlight the topic were identified. The search queries that set up the search strategies have been created by the combination of the related keywords ”ransomware,” ”backup,” ”data protection,” ”cyber recovery,” and ”challenge.” The keywords are selected with regards to answering the literature review purpose.

The inclusion and exclusion criteria for the search were:

- Inclusions

  - The research notes ransomware attack patterns toward backups
  - The research talks about data/system backups in the context of ransomware or disaster recovery. This could be a backup strategy, a solution, or an algorithm. Challenges encountered during recovery are also included.
  - The research was published between 2019-2023 (The last five years). This criterion is justifiable as the thesis considers the most up-to-date knowledge and is not interested in the topic's history. Cybersecurity is a very active and ever-changing field.

- Exclusions

  - If the detection algorithms (or solutions) in the paper are not explicitly related to protecting backups, they have been excluded.

An overview of the included and excluded reports regarding quantity and reasoning can be found in Figure 2.1. The first screening looked at the title, abstract and included a quick screening of the full text of the found 220 documents. After the first screening of the searches, 90 documents were found valid for further analysis. Among these 90 documents, 41 have been found relevant, and 48 have been excluded. The 41 papers have been reinforced by a previous search that consisted of all reviews in the ransomware field, increasing the number of prior research analyzed. Further analysis has been done on texts found by utilizing backward snowballing from relevant texts to create a comprehensive literature review. The final number of texts relate to reports that have been found relevant and fit for inclusion in the thesis. Thus, the literature review has been conducted with rigor in order to understand the research field.

2.2. Literature Findings About Backups

Scoped on backups, one can find ample information in the literature. The importance of backups in organizations is consistently highlighted. [33] states that restoration from a backup will allow any Android device to be saved from the ransomware impacts after a factory reset. [34] notes that robust backup strategies will eliminate most of the threats that ransomware pose. [35] argues that if there are backups of the compromised device available, the attacker will have a tougher time in order to get control of the network. [36] adds that backups reduce the costs and the time to remediate traces of the attack and keep the company's reputation high. [37] states that even if a ransom is paid and all data are given back, the system would still remain compromised if there are not sufficient backups allowing for system recovery. According to [38], there are always limits to the recovery of data after a ransomware infection if there are no backups. [39] shows that willingness to pay is very dependent on the existence of backups. All of these examples show that in the literature, the necessity of backups is well established and that the entirety of the recovery efforts is based upon their existence. However, [40] notes, although there is ample evidence in favor of backups against ransomware, only a small portion of companies have regular backups in place. Considering the evidence, it makes sense that Ransomware would try to circumvent the protection of backups and try to destroy them in order to secure their position for a ransom to be paid.

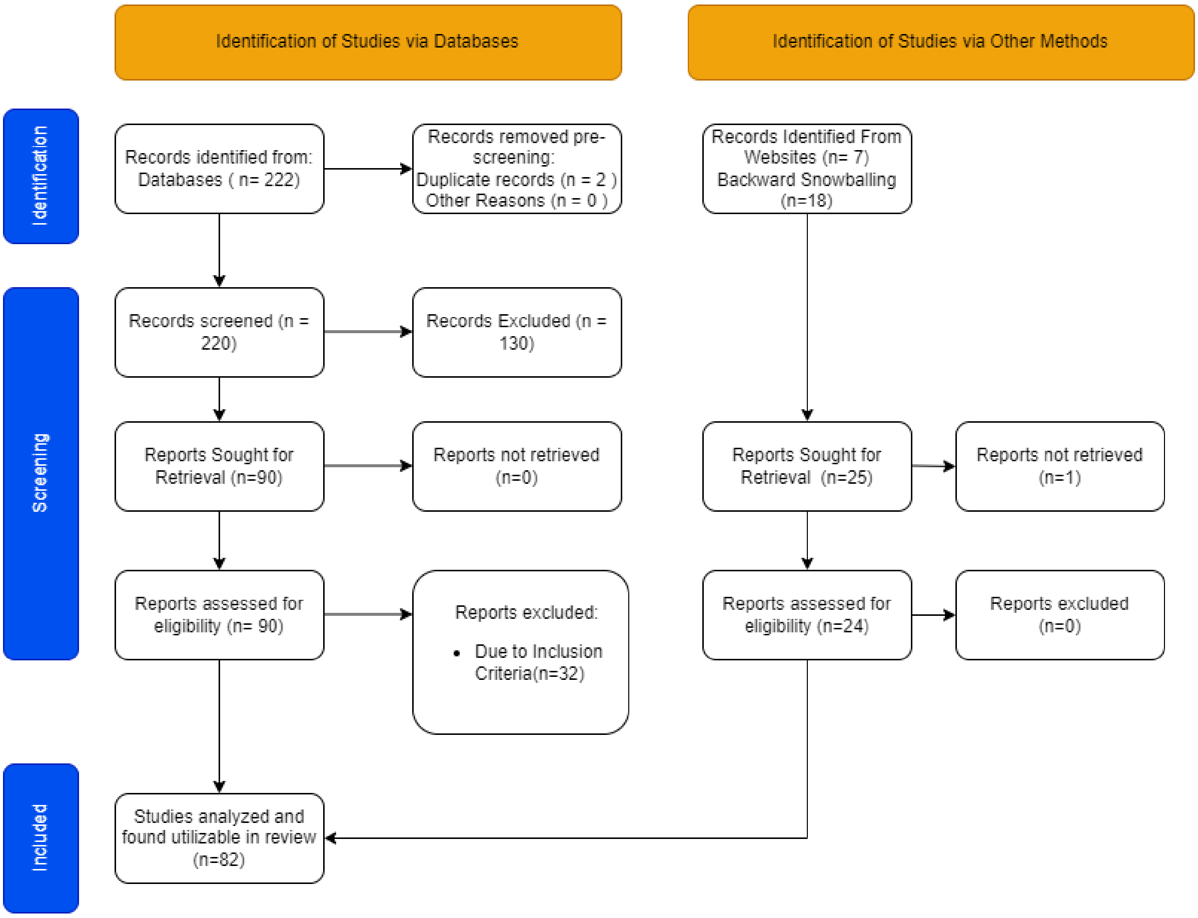

Identification of Studies via Databases
Identification of Studies via Other Methods
Identification
Screening
Included
Records identified from: Databases (n= 222)
Records removed pre-screening: Duplicate records (n = 2) Other Reasons (n = 0)
Records Identified From Websites (n= 7) Backward Snowballing (n=18)
Records screened (n = 220)
Records Excluded (n = 130)
Reports Sought for Retrieval (n=90)
Reports not retrieved (n=0)
Reports Sought for Retrieval (n=25)
Reports not retrieved (n=1)
Reports assessed for eligibility (n= 90)
Reports excluded:
Due to Inclusion Criteria(n=32)
Reports assessed for eligibility (n=24)
Reports excluded (n=0)
Studies analyzed and found utilizable in review (n=82)

Figure 2.1: Prisma Flow Diagram (Adapted from [32])

2.2.1. Ransomware Attack Patterns on Backups

This section focuses on the ransomware attacks that specifically target backup data so that recovery is impossible. Ransomware attacks on backups are noted in literature through mainly noticing that certain services are vulnerable to being exploited. These exploitations can be automatic, or more manually orchestrated. An overview of these ransomware attacks is given in Table 2.1. This section reveals findings that are important to be able to respond to the first subquestion, namely, ”How do the most salient ransomware families target and compromise backups?”

Table 2.1: Ransomware Attacks

Attack	Mentioned In
Shadow Copy Deletion	[6, 11, 12, 41]
Exploiting Vulnerabilities & Evading Countermeasures through Advanced Persistent Threat (APT) Attacks	[9, 12, 34, 39, 42, 43, 44]

The overview shows two main categories of information being present in the literature. Although not too in-depth, it is known that ransomware try destroying the backup copies, exploit vulnerabilities and evade countermeasures through Advanced Persistent Threat (APT) methods to delete backup files.

Shadow Copy Deletion

One of the exploitation procedures of ransomware is to utilize the Volume Shadow Copy service of the Windows OS to delete volume shadow copies (system snapshots). There is ample evidence in the literature (as noted in [6, 11, 12, 41]) that this indeed is a main and a well-documented target for ransomware. The Volume Shadow Copy Service (also known as Volume Snapshot Service) of Windows is a technology that has been introduced since Windows Server 2003. Its main use is to backup application data without closing down the applications and to backup large data segment by segment. It creates snapshots of the OS. Snapshots are the state of the system at a particular point in time. They have differences from backups mainly regarding where they are stored (in the same location as original data), and how long they are stored (shorter term). The Shadow Copies can be utilized to recover the OS back to previous recovery points through the system restore function. However, [45] notes that snapshots are less secure than backups. [46] notes that Ransomware attacks the volume shadow copy service (named, vssadmin.exe under processes) simply by utilizing the following code from a privileged account ”delete shadows /all /quiet” They also note that there are more ways to delete the shadow copies on the Windows OS. [46] gives the example of utilizing PowerShell, an interactive command-line application, or other services such as ”wbadmin” or ”wmic”. However, they do not delve into how these attacks are performed as they do with vssadmin.exe.

Exploiting Vulnerabilities, Advanced Persistent Threat Attacks

Ransomware developers consistently try to find processes to exploit in other OS and devices. For example, [43] notes that SSD solutions that were previously employed against ransomware are currently exploited by ransomware through three new methods. First is that the garbage collector (GC), which deletes the invalidated files in the SSD after new file write operations, is exploited by consistently writing to the SSD and forcing the SSD to release the backup data. Second, ransomware does a timing attack that slows down the speed at which data is encrypted and thus circumvents detection of the device-level protection. The third way of exploitation is to utilize the trim command available in SSDs to physically erase the backup data.

It is important to note that current Ransomware attacks tend to be operated by Advanced Persistent Threats (APTs) [9] who do more targeted attacks against high-value organizations [34]. These attackers may manually move through the system and detect, infect, or alter backup files. If they cannot manipulate the files, they tend to attack the systems that create them. These attacks are not limited to services described before that belong to the system OS. For example, an attacker who identifies that a victim uses tapes to store their backups may attack the main system that overviews those tapes. However, this is a very complex topic, as malicious attackers can change their procedures to fit the environment of the victim.

2.2.2. Backup Strategies

Considering the ransomware threat, and their increasing attacks that target backups, it is essential that organizations adapt their operations to make room for correct backup strategies. Backup strategies relate to practices that organizations can employ in order to ensure backup integrity and enable faster recovery. Strategies in the context of this thesis, however, do not relate to applicable solutions such as Amazon Web Services, algorithms, etc. These are noted seperately under Section 2.2.3.

Literature on ransomware highlights a plethora of backup strategies. It is important to note that many backup & recovery strategies do interact with one another. A simple example is that creating offsite backup storage may also lead to backups being on another network with different authentication and access control protocols and thus being both physically and virtually inaccessible from the main network. This would provide mitigation for both disasters (like fire, etc.) and cyber crises. However, to provide for a basic overview, the most mentioned strategies are given in Table 2.2. The following paragraphs describe strategies relating to the generation, preservation, and testing of the backups. These findings of this section and Section 2.2.3 partially answer the second subquestion, namely, ”What are the current state-of-the-art backup & recovery practices that are (or can be) employed by companies to safeguard their data and ensure operational continuity?”

Table 2.2: Backup Strategies

Strategy	Mentioned In
Regular (Periodic Backups)	[7, 9, 19, 22, 39, 42, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57]
Securing Backups, Access Controls and Authorization	[12, 19, 33, 50, 55, 57, 58, 59, 60, 61]
Offline, Offsite and Redundant Backups	[35, 36, 37, 42, 47, 48, 52, 53, 55, 56, 57, 58, 60, 61, 62, 63, 64, 65]
Proper Backup \ Recovery, Incident Response Plans	[8, 11, 23, 35, 48, 56, 57, 59, 62, 66, 67]
Monitoring, Validating and Testing Backups	[11, 35, 56, 61, 66]

The table shows a balanced attribution to the strategies. In a quick glance it shows that what is most essential is to have regular (periodic backups) that are well secured, is offline offsite and redundant. In an organizational level, backup & recovery must also be supplemented with proper documentation and their corresponding incident response plans. The monitoring, validating and the testing of the backups is also essential.

Regular (Periodic Backups)

Given the table, creating regular (periodic) backups are pretty self-explanatory. Ensuring this will decrease the impact of the ransomware attack. Organizations must ensure that their data and systems that are deemed essential should be regularly backed up. [41] goes one step ahead here and recommends that due to healthcare patient data being critical, daily or other regular backups might prove insufficient and that organizations must try to look for a continual backup approach. [68] agrees with this automatically maintained backup approach. Although a sensible strategy, the amount of data nowadays that is generated would create a challenge for organizations to backup. [59] recommends a thus minor shift in the regular backup strategy. The backups could be created with regard to the organizational value of the data. This implies that important data can be backed up more frequently, whereas less critical data will not. [66] gives support to this as they note that it is important to see and establish what sources of data must be safeguarded and backed up. Thus, not all data is necessary to be backed up.

Securing Backups

Having regular backups is only one side of the coin. Once the backups are created, they must also be secured. They can be secured through different means. [56] gives a machine learning-based detection method that can identify if a backup is altered or not. Monitoring backups through these detection methods can prove critical as, if ransomware infiltrates backups to hide or create garbage, then there is no purpose in recovering from them.

To prevent attacks on backups, proper access control and authorization in the systems that hold the backups is another important aspect. [19] notes that the backup solution must ensure that the snapshots [or other backup files] must be inaccessible from user hosts. [58] gives an example of the Windows-based utility named ”Controlled Folder Access,” where only trusted applications can access designated folders. This utility can be found in Windows Security, and an admin account can identify and modify the list of trusted applications. If an application is not trusted, they do not get access and cannot alter any files in the folder. [55] proposes a novel ransomware protection scheme (which they base on the blockchain) and suggests that only authorized users should pass through the gateway and upload or read backups. [59] notes that there must be proper restrictions in place and that backups must not be accessible through hosts. [60] notes that ensuring proper access controls can prevent ransomware from accessing files and directories. [57] argues that strong passwords must be utilized for access control. Strong passwords will then relate to stronger security on accounts with privileges to access the backups. [12] gives insight into the Active Directory, which is Microsoft’s user directory service that is ”arguably the most popular solution for organizations to manage and organize their staff’s IT profiles for authentication, authorization, and accounting purposes.” They note that this domain controller (Active Directory) must be maintained properly, and correct policies must be applied to correct user accounts. [61], on the other hand, focuses on network access and advises, ”No port access is allowed to the backups, except specified ports needed to run the backup software, and only internal connections should be allowed.” Although the literature sometimes focuses on a specific application of access controls, the main idea is the same. No matter what, organizations must secure access to the backups.

Another security measure could be to ensure that proper patches and upgrades are applied to the systems that contain the backups [12, 50]. Unsupported versions of software, etc., are targets for exploitation. Archive servers that are corrupted by ransomware should also be properly restored [22]. Shadow copies on the other hand can be secured by ensuring that unused services such as the ”vssvc.exe” or the Windows Script Host services ”wscript.exe” and ”cscript.exe” could be disabled [9, 52]. Of course, here, there is the assumption in both references that the organization does not need these utilities. The security of the backups could be expanded by utilizing encryption. Encrypting backups will ensure that they cannot be accessed or altered by attackers. [47] notes by citing [69], that this is feasible if a one-time pad encryption algorithm is utilized. This way, the encryption can occur multiple times, whereas the decryption is done only if the need arises due to an attack.

Organizations are advised in the literature to also consider practices such as creating safe zones, ensuring proper network segmentation and data segmentation, backup data mirroring, or creating air-gapped immutable backup copies. The safe zone can be described as a file system region where ransomware is very less likely to access by design. For example, [70] created an application that keeps all the files of a user by compressing them. This was called the Safe Zone, and it was kept in a non-stop write mode which made it so that no user could alter the files. The application also logs and tracks modifications made in the parent folders of the files added to the Safe Zone. [58] and [47] note that the Safe Zone application could be used for backups against ransomware. The latest backups whose integrity is in place and kept in the Safe Zone could be accessed by the organization and used for recovery. [49] proposes another backup system. Here they create 3 zones named the unsafe, middle, and safe zones. The unsafe zone is mounted to the local machine and is prone to ransomware attacks, whereas the safe zone is isolated from the local machine as well as the remote computer system over the internet. Their system works by a docking and undocking notion where if backups need to be made, the unsafe zone and the middle zone are connected, and the data are backed up. Afterward, the connection is severed. When the organization is sure that there is no infection in place in the middle zone and file integrity is proper, they then basically mount the safe zone and middle zone together to hide the backups.

Ensuring proper network segmentation is discussed by [71] and [61]. Network Segmentation relates to dividing the network into a main domain and subdomains that will provide layers of security. The main network can accept credentials from the subdomain delegated specifically for backups with a ”one-way trust.” The subdomain, however, will not accept credentials from the main domain. There will be different access protocols and policies within the subdomain so that ransomware will have a very difficult time even if it can spill into the subdomain. [72] notes that this segmentation strategy could also be utilized in data. For example, in the Fort Collins Loveland Water District case in 2019, a failure reason for a ransomware attack was that customer payments were handled by a third-party vendor who had sensitive data, and thus, nothing of importance was compromised. Thus, segregating sensitive customer and daily operation data might be useful [72].

The method of backup data mirroring is noted by [73]. It is a mechanism they propose to employ so that electronic medical record systems can be frequently updated for backup purposes. Data mirroring, or a mirror backup in this context, is the automatic exact replication of current data to a specified secondary site where the backups are kept. [74] notes that this mirroring is used in sectors that have 24/7 upkeep necessity, for example, banks, and could provide a redundancy mechanism. If one disk fails, the organization can shift to the second site. However, although [73] does not mention it, there are challenges with mirroring; for example, if there is a corruption due to an improper shutdown, etc., in the main database, the mirrored database will also be corrupted. [74] and [75] note that data mirroring cannot replace proper offsite backups.

Another security solution, very similar to network segmentation, which is deemed state-of-the-art, is air-gapped immutable copies [54, 66]. This is a newer and trendy approach to data backup and storage that focuses on ensuring proper security and cyber resilience in organizations. The air gap means that there is a physical or logical isolation between two systems or networks. Physical isolation could be isolation from the network or a literal separation of the backup storage. A logical isolation, on the other hand, are barriers created with methods such as the previously described network segmentation. An air-gapped environment, in this context, means that the storage that contains the backup is disconnected from the network or the internet. Immutability, on the other hand, means that backup data cannot be modified or altered. The backup data becomes read-only, also noted as write-once-read-many (WORM), and thus cannot be changed. This makes it harder for attackers to alter backup copies.

Offsite and Offline Redundant Backups

There is also a high amount of literature that describes explicitly having offsite and offline backup copies. This is similar to prior described air-gapped immutable copies however are not necessarily immutable. They are, however, similar in having a plethora of physical and logical controls to establish security. [64] argues that the best-known backup method is 3-2-1 method. The 3-2-1 method consists of creating three copies of data, utilizing two different storage media, and having at least one offsite backup. Three copies relate to creating redundancy with additional copies, having two different storage media or devices (incl. external devices [42]) diversifies backup locations and eliminates the single point failure risk. The offsite backup will then provide protection against disasters as well as cyber-attacks. [62] agrees with this as they note that geographically distributed redundant data backups will allow swift recovery with minimal disruption. Other references that are provided in Table 2.2, also note similar arguments for offline offsite and redundant backups.

Generating backups and ensuring their security is good; however, this does not relate to having recoverable backups during a crisis. Backups must be regularly tested from end-to-end to see if recovery is possible and that they are fit for purpose. Backup integrity is not only dependent on a virus attack but could simply be a network failure during the backup process. The testing must be done and backups monitored in order to proactively solve any problems that might occur prior to a crisis.

Monitoring, Validating and Testing Backups

A last backup strategy that is many times implied and sometimes stated outright is that there must be backup and recovery plans that are properly documented with established roles and responsibilities [48] that encompass everything that has been described within this section and more. [62] notes that these plans must entail ”Critical Business Functions (CBFs), Maximum Acceptable Outage (MAO), Recovery Time Objectives (RTOs).” Here an addition could be to have Recovery Point Objectives (RPOs), as although elapsed time to full recovery is important, to what point the system recovers to is also essential. [66] implies that these plans (and policies) should also entail the backup types, frequency of backups and testing, and storage media to be used. These backup and recovery plans are to be guided by organizational objectives and to guide policy within the organization to ensure cyber resilience against attacks.

Backup strategies are generally identified as controls in international cybersecurity standards that are set by established entities. These cybersecurity standards, also called cybersecurity frameworks, have not been so visible according to the author's searches. The only two references found regarding cybersecurity frameworks were the NIST Framework [76] and the Critical Information Security (CIS) Controls [72]. [76] goes into creating a serious game to bring awareness to the NIST Framework, whereas [72] applies the CIS controls to the cybersecurity incidents in the water sector.

2.2.3. Backup Solutions

There are a number of solutions that offer backup services and ransomware protection that are mentioned in the literature. An overview is given in Table 2.3

Table 2.3: Backup Solutions

Solution	**Mentioned In**
Software Based Solutions	[43, 59, 61]
Hardware Based Solutions (Including NAS)	[42, 43, 49, 61, 64, 77, 78]
Cloud-based Solutions	[6, 19, 49, 60, 61, 62, 64, 67, 71, 79, 80]
Detection and Recovery Based Solutions	[51, 68]
Blockchain	[55, 80]

The table includes 5 main categories for the type of solutions that are prevalent in the backup & recovery domain that can be utilized by organizations. What is important to note is that although Detection and Recovery based solutions are written explicitly in the table, both software, cloud or hardware based solutions are able to utilize detection and recovery. However, it is important to note it explicitly as not all software is designed to promote this type of utility. Therefore detection and recovery does not have its own section in the further subsections but is spread across. Another important thing to note is that the highest number of solutions that is described in the literature is cloud-based solutions. A probable reason is because it is easy to start using, to scale, and makes things easier regarding backup and recovery by providing an offsite alternative out-of-the-box. The following subsections give an understanding of each solution type.

Software Based Solutions

Regarding software tools, [59] gives an example of two anti-ransomware tools belonging to Kaspersky and Malwarebytes. The Kaspersky tool employs a system watcher technology where they monitor file usage patterns and the network communication of processes. Their behavior log is then compared to the signatures with Kaspersky's latest threat lists, and the process is terminated if it matches ransomware. During this time, the tool keeps backups of all the files that are being modified by the observed process, and if the process is found malicious, the backups are then restored automatically. The Malwarebytes tool also works in a similar way and monitors the file access pattern of a target process. It can also detect bulk encryption. The restoration and backup are done similarly to the Kaspersky tool. For a backup solution, [61] notes the VEEAM software for storage purposes.

Hardware Based Solutions (Including NAS)

[64] and [49] note the usage of NAS storage in order to save backup copies. [49], however additionally mentions that NAS storage is also not fully safe for ransomware as they do not isolate the backup systems automatically and can be accessed if not secured properly. Synology, and QNAP are noted as two NAS storage providers that could be used as offsite backup solutions. Regarding offsite backup solutions [61] notes that Iron Mountain is an alternative physical offsite storage vendor.

[42] lists some more device-level solutions in the fight against ransomware. They note that solid-state drives' flash memory characteristics allow protection without additional backup devices. In a simple way, their explanation delves into how SSDs write information in a new location and invalidate the previous data that was already written, and then these are deleted through a mechanism called garbage collection (GC). It is as if a file's altered version gets copied as the new file while the previous one is set up to be deleted. These invalidated versions, then, could actually be utilized as backups as these invalidated files are the previous versions. [42] lists that FlashGuard and SSD-Insider are the most prominent solutions that utilize this functionality to create backups. They note that these solutions detect ransomware through its read-and-write patterns in real-time. Ransomware typically reads, encrypts, and rewrites the files in a bursty manner to the storage. [42] also notes AMOEBA as a contender to the previous two as it can also detect non-bursty ransomware attacks and optimizes the backup storage occupied by maintaining only necessary backup copies. [78] promotes their own SSD solution named the RansomBlocker (RBlocker), which works in a similar mechanism but is promoted as a low-overhead and ransomware-proof solution compared to others. Lastly, [43] also promotes the RSSD, which is their version of an SSD that is ransomware-proof against the newer types of ransomware attacks towards the SSDs.

[77] then, proposes KEY-SSD, which is also a device-level solution against ransomware attacks but is a bit different than previous proposed SSD solutions. It is a disk drive that uses an access-control mechanism to alleviate data breaches. They note that even if ransomware has admin privileges and can access the file system, it will be blocked by the disk-level access control due to not having the disk's registered block key.

Cloud-based Solutions

After the software and device-level backup (& security) solutions, cloud storage is seen as a relatively cheaper solution utilized by organizations for storage purposes (incl. backups). [62] notes that cloud computing can be a lifesaver for especially middle to small organizations for generating and maintaining backups. The benefit of utilizing such a cloud system is that the networking and storage capacities are basically subcontracted to a third party, creating a cost-effective alternative. With the cloud, organizations are able to significantly expand their infrastructure resources. [62] however also notes the disadvantages of utilizing the cloud. Customers will not be able to control where their data will be stored, or they may have security and data confidentiality concerns due to third-party involvement. This way, if the cloud customer is dealing with sensitive data, they might find it difficult to comply with corresponding regulations of data management. [67] gives a cloud service reliability evaluation framework in order to assess the accessibility, continuity, data backup, and recover-ability of cloud providers to alleviate customer concerns. To alleviate the challenges, [64] reveals a distinction between private and public clouds. Public clouds are owned by third-party organizations such as Google, Amazon (Amazon Web Services), and Microsoft (Azure) [61]. Private clouds on the other hand are hosted within an organization. [64] argues that the private cloud infrastructure may overcome the challenges of using a cloud-based system and prove to be more cost-effective in the long run.

[6] lists two other cloud-based applications against ransomware. CLDSafe is noted as a cloud-based file backup system that is designed to backup both files and shadow copies of the user's system. It does this through a method of checking a similarity score between the new and the old file. It does incremental backups, and the application is able to detect infected backup files through a hash technique. Another application is noted as CloudRPS. This application collects data from the cloud in order to detect and analyze abnormal behaviors in the server, files, and network. A limitation however is that syncronization takes time and is limited in terms of storage space.

Blockchain

As another measure that is less seen in the literature, [80] notes that for secure backup data blockchain cloud infrastructure could be utilized for authenticity and access control. [55] on the other hand creates a novel data backup scheme that is based on blockchain. However, neither methods could be utilized to handle the sheer amount of data and computational loads effectively.

2.2.4. Challenges in Backups & Recovery

There are a lot of recorded challenges with respect to the adoption of proper backup procedures and plans and difficulties during recovery from backups. The biggest concern factor regarding challenges is found to be regarding backup integrity. An overview of all the challenges faced during recovery can be found in Table 2.4

Table 2.4: Backup & Recovery Challenges

Challenge	Mentioned In
Backup Integrity	[7, 9, 12, 19, 37, 42, 44, 46, 50, 52, 53, 54, 55, 58, 59, 61, 63, 64, 79, 81, 82, 83, 84]
Higher Resource Cost	[12, 19, 22, 37, 68, 78, 83, 85, 86]
Absence of Regular Backups	[47, 87]
Low Cybersecurity Education and Awareness	[12, 57]
Unpatched Vulnerabilities in the Backup System	[12, 87]

From the overview table it could be seen that the main perceived challenges by the academic literature, are that organizations face loss of backup integrity which makes them unable to recover from backups in the time of a crisis and that the higher resource costs make them ineligible to be able to scale their backup & recovery practices to a higher maturity level. The following paragraphs go more in-depth of each challenge.

Starting with backup integrity, [7, 44, 53, 64] note that in many cases the backups are encrypted by the attackers, and with no possibility of decryption, the backups cannot be utilized. [54, 61, 84] look at this from the side of corruption of the backups. Many attackers choose to corrupt the backups so that, although they look like everything is normal, they are unusable during recovery.

[44, 46, 55, 58, 63, 79, 82] note shadow copies and backups also cannot be relied on during recovery if they are not properly secured, monitored and tested. This is due to attackers primarily targeting these systems. [55] additionally notes that this also is a prevalent concern even if cloud solutions are utilized as it is still easy to access, tamper and delete backup files. [49] adds that even if cloud servers and external storage such as Network-Attached Storage (NAS) are utilized, they are still not safe as malware tend to find ways to spill over to them through the internet. As long as the backup storage is weak, ransomware will infiltrate and destroy it [59].

[37] notes that within their examples, they have seen that the recovery period increases due to improper backup procedures. [87] and [47] identify that there are absence of regular backups in organizations. Not having regular (and frequent) backups leads to having gaps in the data after recovery as [9] notes that if the backup is not recent, most changes made after the last backup will not be available. [23] notes that in healthcare facilities they have found that many hospitals did not even have defined strategies or plans, and thus backups, to alleviate ransomware intrusions.

[22] argues that there is also a lack of relevant professional information regarding cybersecurity. They note that the adoption of a backup system is often discussed but mostly left up in the air, and no action gets done. This could be caused by lower cybersecurity education and awareness. [12] notes that even with the publicized impacts of ransomware, organizations are falling behind in taking notes about lessons learned from attacks. Many organizations even tend to have unpatched systems which are vulnerable to exploits used years before [87]. To this end, [57] raises questions about the effectiveness of cybersecurity education, awareness, and training interventions in organizations.

Backups also tend to cost higher resources. [19] notes that backup maintenance takes up a lot of CPU and disk usage. These backups also take up space in the local disk. [68] and [83] note that there is a lot of data bloat and excessive disk usage due to not having a consensus on what backup strategies or backup solutions. [78] argues that overhead caused due to backup systems may degrade the user experience.

[85] and [37] with their observations argue, that due to the relative cost of decryption or paying for the ransom versus establishing backup systems for recovery, creating backups may not be always beneficial. Also related to this, many organizations however either lack the proper resources or the allocation of the resources within the organization. [22] argues that small and medium enterprises have capital scarcity and insufficient human resources. This then leads to these organizations not being able to build their own Information Security (IS) protection systems. They also argue that most small and medium-sized organizations tend to see IS as ”an option” and not a requirement. [12] states that organizations often fail not only in patching vulnerabilities but allocating sufficient resources toward IT and cybersecurity that would help during recovery.

2.3. Research Gaps

The research gaps that are distilled from this literature review are that, firstly, ransomware attacks on backup and recovery are not well explored. There is a knowledge gap on how the ransomware currently targets backups and recovery. At the moment the literature either considers data all together (backups and primary data) and explains the main behavior of the ransomware [9, 34, 42, 44] such as infecting and destroying data. On the other hand, the literature repeatedly identifies a specific attack such as deleting shadow backup copies through windows utilities, but does not give extensive information on how these attacks are commenced in operating systems [6, 39, 41]. Although there is indeed some information on how ransomware use SSD specific hardware mechanisms to destroy backups [43], this is limited to only SSD hardware. Moreover, there is no specific information on how and which vulnerabilities ransomware has used on applications that govern backup processes in the latest years or how ransomware circumvent end-point or backup software defences. Then it becomes essential to explore and identify how ransomware is known by the most credible antivirus and backup software vendors with respect to their attacks towards backup & recovery.

Secondly, based on this literature review, there is a knowledge gap with regards to what do the well-established cybersecurity frameworks, where some are regulations to be complied to by organizations, offer as best practices in the context of backups & recovery. Mostly, the cybersecurity frameworks are only noted as ”state-of-the-art” and their names are sometimes given and sometimes not, but no detailed instructions from them are identified in the literature [38, 50, 79, 81, 88, 89]). Exceptions to this however are the NIST Framework, and the CIS controls. Each of the aspects of the NIST Framework are operationalized in the work of Onwubiko, [48]. This is also true in the work of Scherb et al. [76] as they create a serious game based on the NIST Framework for Cybersecurity. The CIS Controls on the other hand are mentioned rigorously in Hassanzadeh’s work [72]. Now, here it is important to note that both frameworks have received updates after Onwubiko and Hassanzadeh’s work (excluding Scherb et al., as they are up to date in 2023). Even without the fact of closing the gap between the previous and the updated versions of the NIST and CIS Frameworks, there is still research necessary into other relevant cybersecurity frameworks such as the IT-Grundschutz [90] (German Federal Office for Information Security) or the ISO 27000 and ISO27001 [91], or the new Digital Operational Resilience Act (DORA) [92] regulations of the EU, etc. [48] cites in their work that ”although there are existing federal policies, standards, and guidelines on cyber event handling, none of them focuses solely on improving cybersecurity recovery capabilities, and the fundamental information is not captured in a single document.” This thesis then provides a capturing device for the strategies highlighted for backups and recovery in these frameworks in order to allow organizations and academia to have a sufficient insight into the comprehensive state-of-the-art. Therefore the main knowledge gap here is, what are the best practices that are recommended by the well-established cybersecurity frameworks in the context of backups and recovery? This, in the light of Onwubiko’s quote from prior, is then expanded because there is also no established comprehensive and systematic overview in the literature of all the best practices regarding backup & recovery against ransomware attacks (or other cyber attacks).

Thirdly, the question ”why do organizations still fall victim to ransomware attacks and have challenges during recovery?” must be explored. [48] notes that ”despite the importance, very little contribution exists on cyber recovery.” The academic literature gives main insights on challenges regarding the backup data integrity and higher resource costs for organizations. However, other than Kumar’s work (2022) [93], there is a lack of knowledge on what makes it actually difficult for organizations to adopt good backup & recovery practices.

2.4. Conclusion

This chapter has belonged to the method for semi-systematic literature review. The literature review has considered 82 academic articles that have been selected from the databases Scopus and ArXiv, by utilizing relevant keywords and establishing an inclusion/exclusion criteria. The chapter has given the research gaps that has served as the basis for the creation of the research question, and its subquestions respectively. The method itself also allowed discovery and overview into what are the known ransomware attacks towards backups, the best backup & recovery practices and challenges that organizations face. The overview over these three focus points have allowed for the establishment of the implications for organizations in the context of increasing backup & recovery cyber resilience. The main ransomware attacks that have been discovered were: shadow copy deletion, exploiting vulnerabilities, and evading countermeasures through APT attacks. The implication of these are that organizations should make sure to only allow the usage of such utilities, that can access, modify and delete shadow copies in the operating systems, in the utmost necessity for only a select few employees. The ransomware can utilize hardware-level mechanisms to be able to destroy data on SSD drives. The implication here is that the organizations must be vigilant that the legit and useful mechanisms in their hardware can be utilized in order to destroy backup data. Ransomware can exploit these mechanisms through APT attacks, as these are coordinated attacks by actors who actively strategize and proceed the attack with respect to the organization at hand. The only way to ensure that the APT attacks do not succeed could be to ensure cyber resilience through best practices.

The main backup practices discovered throughout the semi-systematic literature review were: ensuring regular backups, securing the generated backups, establishing proper backup & recovery procedures, and monitoring and testing backups. The Organizations can ensure regular scheduled backups through also utilizing a continual backup approach if the data criticality is too high that nearly no data loss is permissible. Categorizing data and establishing backup measures with respect to it is key to not face increased data sizes. The security of the backups must be established through dimensions such as proper access control and authorization. Strong passwords must be utilized, and patch management must be correctly monitored. As described before, access to utilities that can destroy backups must be limited. The backups could be encrypted for ensuring that even if the data falls into the hands of the attackers, it is not usable. Moreover organizations should create safe zones for their backup data, create airgapped environments for their backups, and ensure network and data segmentation. Creating offline and offsite backups will relate to higher cyber resilience of the backups against attacks. The monitoring, validation and the testing of the backups then ensure that they are fit-for-use in time of a crisis.

Organizations could also utilize different backup solutions to make the processes of backup & recovery easier. These solutions could be software-based, hardware-based, cloud-based and, although quite seldom, blockchain-based. These solutions sometimes offer detection and automatic recovery options which could be a good choice for organizations to invest in, in order to increase their cyber resilience.

The main challenges that have been described in literature are: problems with backup data integrity, higher resource costs for organizations, the absence of regular backups, low cybersecurity education and awareness and unpatched vulnerabilities. The main implication for recommendations here is that the backup data integrity is usually lost as attackers get access to storages that the backup data is kept in. In order to keep this to a minimum organizations must ensure that they utilize the previously described best practices for backups in order to keep them secure and have redundant backups in seperate places to be able to recover their systems. The higher costs on the other hand are generally caused by the cpu, disk usage, and the overhead of the maintenance of the backup data. However many organizations either lack proper resources or the allocation of resources to implement strong cyber resilience practices. A recommendation on this side would be that organizations must adopt a ”cybersecurity is a necessity” mindset rather than perceiving it as an option.

3 Qualitative Content Analysis

This chapter is going to highlight information that have been found among 102 sources (on top of the academical documents from prior) that relate to ransomware attacks toward backups, backup & recovery practices, and challenges. Prior to listing the findings, the procedure that was applied and the description for qualitative content analysis is given in this chapter. The procedure and the description contains information on what data was collected and how it was analyzed. It also identifies how the method will tackle the aims, objectives and the research questions in the thesis. The following sections after that all begin with an introduction, continued by the presentation of their inductively generated codebooks. The codebooks serve as main findings that answer the first three subquestions, and by correspondence create the basis of answering the main research question, namely: *”What recommendations can be made for organizations to help them develop good backup and recovery strategies against the growing ransomware threat through a systematic and methodological approach?”*. The answer for this main research question is given in Chapter 5. The codebooks are followed by the abstraction figures where the main categorization of the findings are given. Then there is an in-depth explanation of concepts per section.

3.1. Description of the Method

Content analysis is a systematic and replicable technique that can reduce many texts into fewer content categories that are based on explicit ruling codes. Qualitative content analysis aims to:

> ”attain a condensed and broad description of the phenomenon, and the outcome of the analysis is concepts or categories describing the phenomenon. Usually the purpose of those concepts or categories is to build up a model, conceptual system, conceptual map or categories” [27]

These aims match well with the aims of the thesis as the main research question wants to be able to provide recommendations to organizations regarding what the best backup & recovery strategies are, that is tempered against the current ransomware attacks and the challenges that organizations face. This recommendation list then is a conceptual map. The reader here can visualize a recommendations catalogue. These concepts, such as ”immutable backups” have come together to be mapped under categories in backup strategies, and that with each category new concepts are revealed. All of which relate to answering the main research question and its corresponding sub questions. The content analysis method is also very applicable as it is very flexible to match the needs of the research design. Considering the pragmatic view that was established in Section 1.7.1, this is good news. The qualitative content analysis can allow the researcher to identify and distinguish what is relevant to answer the research questions very quickly. The qualitative content analysis also fits the usage of open ended exploratory questions.

Quantitive content analysis was not utilized because it is more of a deductive approach that is limited by either a prior theoretical or a conceptual framework. It becomes more of a numbers game to establish trends, or to identify quantitive patterns. However it requires very narrow, hypothesis based questions. It further requires a statistical insight to establish its validity. None of which are qualities that is required for this research.

The data sources that have been utilized for the qualitative content analysis are antivirus and backup software/hardware provider reports, and their corresponding security blogs. Moreover statements of victim organizations have also been utilized. This data were used both as primary and secondary sources. The quantitative details, such as how many times in which reports a certain concept were mentioned are primary data that was generated after the content analysis procedure. The contents of the reports however have made up the qualitative part of the thesis and thus is used as secondary data.

The data collection for the qualitative content analysis was done as follows: by utilizing the a purposive sampling strategy, a google search was made utilizing the combination of the keywords ”ransomware”, ”recovery”, ”challenges”, ”backup strategies”. The data then were collected as article pdfs by utilizing the NCapture method of the NViVo qualitative analysis software. The data collection has been stopped when a distinct data saturation was reached. This is when the material starts repeating itself.

The data analysis has been conducted with respect to the guidelines specified in [27]. The data analysis has three main phases, and these are; preperation, organization and reporting. With respect to the context of the thesis, preperation phase starts with selecting the unit of analysis. The unit of analysis could be words, sentences, paragraphs, etc. After a couple of trial runs it was decided that a sentence is best suited to capture the meaning of a mentioned backup strategy, a challenge or a ransomware attack. It is not granular so that the idea flow is too fragmented, but it is not large as a paragraph so that it does not contain the spill of other backup strategies or challenges.

Another decision is made in order to only analyze manifest content. This is justified as the aims of the thesis are specifically to create a recommendation list for organizations. It does not look into finding out about the discourse, the implied meanings behind words, etc. This is also in support of the research philosophy that has been mentioned before as if the perspective was latent content, the underlying philosophy selected for the thesis would have been interpretivism.

In the next stage, the qualitative data has been organized. This has been done through first open coding, then creating categories to establish links between the codes and to destroy duplicates, then to settle on an abstraction. The open coding in this context means that each sentence that had a way to answer one of the three sub questions, and thus the main question, have been coded. Then the codes have been tried to be classified under mutually exclusive categories. However, due to the very interconnected nature of the topic, and the inductive research paradigm this constraint was relaxed. The main thinking during the classification under categories was for example: ”are all airgap backups immutable?” If the answer is yes, then these two concepts may very well belong to the same category. However as the answer is no, and that there is no overarching category, they have been classified in different categories. [27] notes that categorization falls to the researcher to choose in qualitative and inductive content analysis. It is quite dependent on the researcher’s ability to reason and to construct meaning. [27] further notes that there are no guidelines for the data analysis of a qualitative content analysis, and argues that this flexibility is both a strength and a challenge of the method. There simply is no ”one right way”. Therefore the qualitative content analysis tried to fit the aims of the thesis, to provide a recommendation list by creating codebooks that are similar to a ransomware attack, strategy, or a recovery challenges catalogue.

The organization stage is finalized with abstraction. Each section under the findings have their own abstractions presented. These formulate a general description of the corresponding research topic and also to create a narrative structure for cohesive argumentation. Then the analysis results are reported through first identifying the categorizations and the trends in quantitative measures. The quantitative measures also allow for identifying which material was related to which ransomware attack, backup strategy or recovery challenge. This continues with the qualitative descriptions of each category and sub categories to present the insights.

In order to deal with the complexity and challenges of the method, NViVo software was used. The software allowed for data collection and open coding in the pdf files. It further allowed to categorize data. Moreover it keeps track of how many times the same code has been coded, which is an impeccible contribution to the research which can identify what the main considerations in the industry reports are. The number of times of a concept has been coded gives information on the differences or similarities between sources and methods, and gives insight into what is the focus of the documents. It also gives an idea about what are then the main recommendations are, that must be noted. Using the amount of times a concept has been noted as a data point is indeed novelty in the literature, as it helps to establish an overview such as this thesis.

In this chapter, Section 3.2 answers the first subquestion: *”How do most salient ransomware families target and compromise backups?”*. This section is divided into 5 subsections. The 5 subsections explore in-depth: how ransomware behave in order to compromize backups during reconnaissance, lateral movement & privilage escalation; how ransomware uses vulenrabilities on backup software & hardware on different operating systems, how the service utilities are misused in order to destroy backups, and lastly, how antivirus and backup software processes are identified and terminated. All these subsections go in-depth and answer how the most salient ransomware families such as BlackByte target backups. This then provides a critical reflection point to generate recommendations that can specifically remedy these attacks.

Section 3.3 on the other hand focuses on answering the second subquestion: *”What are the current state-of-the-art backup & recovery practices that are (or can be) employed by companies to safeguard their data and ensure operational continuity?”* This section is segmented with respect to the data sources they utilize. The section identifies the difference in importance between a cybersecurity framework which is created by credible and established entities such as the US National Institute of Standards and Technology, or the EU Commission (DORA regulations); and the reports of antivirus or backup software providers. Albeit both data sources having their own contribution with regards to exploring the topic at hand, the cybersecurity frameworks are seen as the core generator for our understanding of what best practices in the field regarding backup and recovery are. Therefore in this section, the best practices from the cybersecurity frameworks are given as an overview among eachother, where similarities and differences on their approach to backup & recovery can be seen. These findings then are expanded by utilizing the reports of antivirus and backup software providers & other material such as cybersecurity blogs. These findings answer the subquestion on the form of a codebook that goes over operational practices & choices for backup storage options wherein all actionable items or policy measures for organizations can be found. The expectation as outcomes in this section are actionable items for organizations to increase their cyber resilience with regards to their backup & recovery practices.

Lastly, Section 3.4 explores the third subquestion: *”What are the main challenges faced by organizations regarding backups & recovery”*. These findings are to be reflected upon in Chapter 5 in order to produce policy recommendations for organizations to answer the main research question. The expected outcome here is that the findings are to generate insight into why organizations might fail or face difficulties in applying the best practices and adopting better cyber security practices with regards to backup & recovery. Therefore the section tries to identify the main pain points, the failures and to delve into understanding why this becomes such a complex topic in organizations in the first place.

3.2. Ransomware Attacks Toward Backups

Ransomware Attacks towards backups primarily happen through doing reconnaissance in order to find the backups, and then doing lateral movement and privilege escalation to access the backups. Moreover, the attackers utilize unpatched vulnerabilities or zero-day vulnerabilities in order to get access to the system. Zero-day vulnerabilities, for example, are found in an update that is rolled out, etc. and serve as a good initial access point. With this access, the attackers will simply delete, corrupt, infect or encrypt the backup files. This is done because the attackers know that destroying the backups will lead to a higher incentive to pay and thus the backups are a primary target. With regard to ransomware attacks, many things are interrelated and everything leads to the aim to obtain more access to the network. More access corresponds always to more damage.

Ransomware can utilize what is called a supply chain attack, where instead of attacking the main IT components of an organization they will attack a third-party component that the organization utilizes to gain access. These third-party components can be breached through unpatched vulnerabilities in their software. The ransomware can then conduct a living-off-the-land attack where they use the utilities that are available in the IT Landscape in the organization to increase their access to more files. Ransomware can also deploy its own tools (where many are actually legitimate) into the network to enumerate and access more of it. Therefore the name of the game for the ransomware actors is, access. [94] notes that with uninterrupted access, the backups will be compromised by the attackers. [95] notes that if the ransomware has access to the internal network, it will be able to achieve domain administrator privileges and thus access to backups.

The codebook that has been generated through the analysis of the literature is given in Table 3.1, which is followed by the analysis of the trends and focus in the domain. The codebook itself represents a summary of the ideas and concepts which answer the first sub-question. The ”#” column on the codebook represents the number of times a particular concept was mentioned, and is cumulative for top-level categories. Explanations of the key insights are made following the abstraction.

Table 3.1: Codebook of Ransomware Attacks on Backups

Category	Sub Category	#	Citations
Backup Integrity Attacks to Storage Management System		8	[90, 96, 97, 98, 99]
	Corrupting Backups	2	[96, 97]
	Delete Backups	2	[98, 99]
	Encrypting Backups	1	[96]
	Infect Backup Copies	1	[90]
Identify and Terminate Services and Processes		7	[98, 100, 101, 102, 103, 104, 105]
Privilege Escalation		10	[90, 94, 95, 96, 106, 107, 108, 109, 110, 111]
Reconnaissance for Backup		6	[107, 108, 109, 110, 112, 113]
Use of Service Utilities in OS		10	[98, 99, 100, 103, 104, 107, 110, 114, 115, 116]
	Boot Configuration Data Editor	1	[99]
	Volume Shadow Copy Service	9	[98, 99, 103, 104, 110, 114, 116]
	Windows Backup Administrator Utility	1	[99]
	Windows Recovery Environment	1	[99]
	Windows Restart Manager	2	[100, 116]
	Windows Management Instrumentation	2	[99, 107]
Use of Vulnerabilities		19	[96, 108, 109, 111, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129]

The codebook clearly shows that ransomware attacks on backups are mainly mentioned through their use of vulnerabilities (19 mentions). This is then followed by their use of service utilities (10 mentions), privilege escalation (10 mentions), backup integrity attacks to storage management systems (8 mentions), termination of processes

(7 mentions) and finally the reconnaissance for backups (6 mentions).

The category of use of service utilities has its corresponding subcategories. The most mentioned service utility here is the volume shadow copy service (9 mentions) and it significantly outweighs the other utilities in terms of mentions made. Other service utilities are the Windows Management Instrumentation (2 mentions), Windows Restart Manager (2 mentions), Boot Configuration Data Editor (1 mention), Windows Backup Administrator Utility (1 mention), and the Windows Recovery Environment (1 mention). Backup integrity attacks on storage management systems category on the other hand is divided into; the corruption (2 mentions), deleting/removing (2 mentions), encrypting, and infecting of the backup copies (1 mention).

The abstraction that has been generated following the analysis of the codebook is given below. The abstraction figure in Figure 3.1 lays out the main categories that have been derived in the creation and further analysis of the codes. The main categories that have been laid out are; the reconnaissance for backups, privilege escalation (and Lateral Movement), the use of vulnerabilities, utilizing of service utilities in the OS, and finally the attacks on backup integrity & storage management systems. The narrative of the further sections thus follows this pattern to delve deeper.

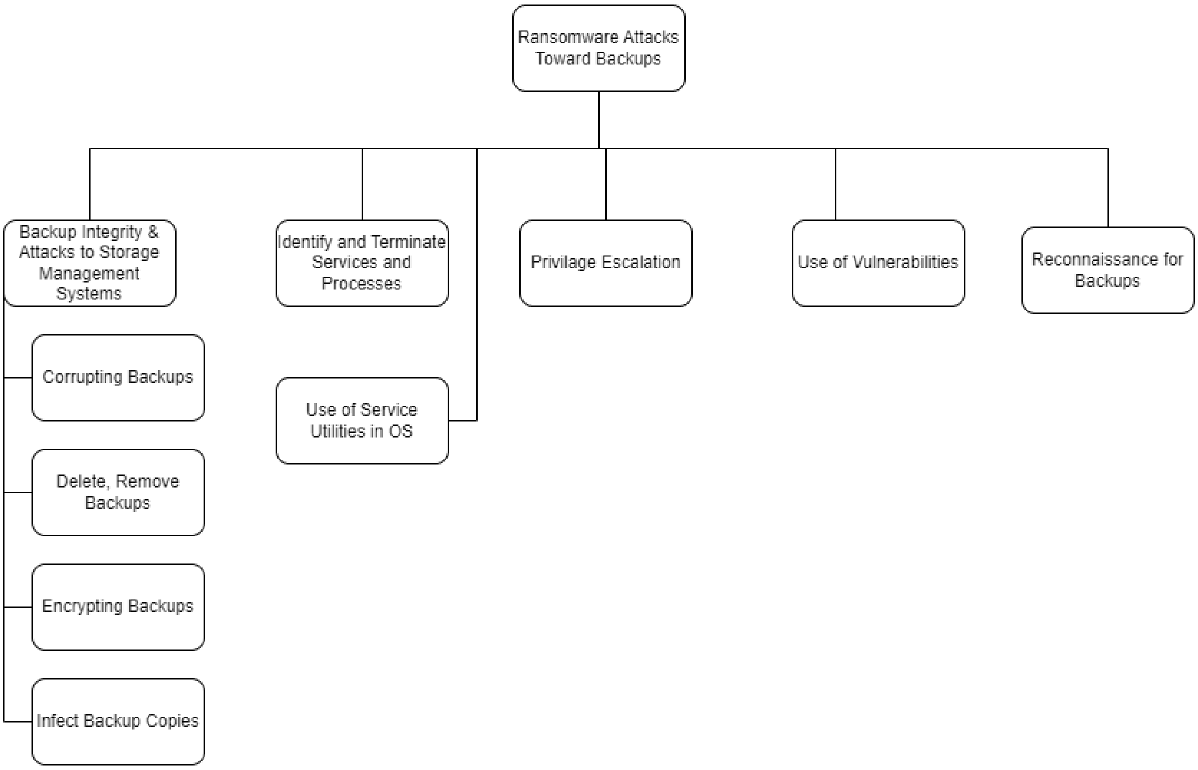

Ransomware Attacks Toward Backups
Backup Integrity & Attacks to Storage Management Systems
Identify and Terminate Services and Processes
Privilage Escalation
Use of Vulnerabilities
Reconnaissance for Backups
Corrupting Backups
Use of Service Utilities in OS
Delete, Remove Backups
Encrypting Backups
Infect Backup Copies

Figure 3.1: Abstraction on the Ransomware Attacks on Backups

3.2.1. Reconnaissance, Lateral Movement & Privilege Escalation

Ransomware actors do reconnaissance by utilizing different tools. This reconnaissance allows them to understand the network that they are dealing with. At this stage, they will be able to figure out what, where, and how's to their attack including destroying the backup copies. The NIST Cybersecurity Framework [96] supports this claim through the example that an attacker may use elevated privileges to compromise backup copies and/or the generation of future backups. The IT Grundschutz Cybersecurity Framework [90] notes that ransomware actors look for network drives that have write access to be able to encrypt data. [110] argues that ransomware finds and encrypts the backups on the file shares. They highlight that many organizations use default directory names, that are published in the documentation of the backup products, to store the backups. With reconnaissance, ransomware finds these files to encrypt them in their directories.

In order to achieve reconnaissance ransomware utilize certain tactics. [107] notes that the Lockbit ransomware uses netscan, and advanced port scanners to enumerate the internal network. They use the AdFind tool to locate the Domain Controller or the Active Directory server. The Domain Controller (DC) is a server in a Windows domain network that responds to security authentication requests such as logins and checking account permissions. It also stores user account information and enforces the pre-determined security policy. The Active Directory (AD) on the other hand is a technology used to manage computers and other devices on a network. The AD service includes information about network resources which includes users, groups, computers, and other objects. [108] notes the tool ADRecon allowed the ALPHV ransomware to gather network, account, and host information in the victim organization's environment. The tool allowed the ransomware, with its high privileges, to generate reports about the AD environment which included the network shares, trust policies, and user account lists. The AvosLocker ransomware [109] used netscan and Advanced IP Scanner to be able to identify network shares in the organization. The ransomware Karakurt [113], on the other hand, used Cobalt Strike to enumerate a network to decide on an attack pattern.

The reconnaissance is often followed by privilege escalation and lateral movement in the network. Privilege escalation happens either vertically or horizontally [96]. Vertical privilege escalation means that a lower-privilege user or application is able to access a higher-privilege user or application. Horizontal privilege escalation is when a user can access other user's content on the same privilege level. In order to achieve privilege escalation ransomware tries to steal credentials [94], utilize brute-force or dictionary attacks [111], or use tools such as Mimikatz, LaZagne, and Nanodump to harvest credentials or use other methods to bypass user account control [107, 108].

To achieve lateral movement [109] notes that the ransomware actors utilize remote desktop protocols (RDP) and the remote management tool named AnyDesk. The lateral movement here refers to the unauthorized movement of attackers or malware within a network from one compromised system to another. The lateral movement is done in conjuncture with privilege escalation prior to impact as was denoted in Figure 1.1.

3.2.2. Use of Vulnerabilities on Backup Software & Hardware

The ransomware can use vulnerabilities to get initial access, increase their privileges, and move laterally throughout the network. On this end, they also utilize the vulnerabilities that the backup software or hardware may have. Thus, it is imperative that organizations do their patch management properly.

There have been multiple examples in the field where ransomware has utilized a vulnerability. AvosLocker had leveraged vulnerabilities in the Veeam Backup and Replication Software [109]. In this particular attack, the attacker was not able to execute their ransomware but could delete Veeam backups through the tool's interface. [130] notes that some of the users are still not up to date with the Veeam software. Even though the databases that were managed by Veeam were encrypted it was possible to bypass it. Veeam, although releasing fixes shortly after the breaches, has also been hit by other ransomware targeting different vulnerabilities [124, 131].

Veeam is not the only software that is attacked. FOX-IT [128] gives an example of attackers leveraging the R1Soft Server Backup Manager Software. They utilized it for both initial access and to control other downstream systems. As backup software usually has high privileges they were able to execute commands on all systems that ran the server backup manager.

QNAP, a supplier of Networked Attached Storage (NAS) devices where backups are possible to be stored, have noted that the Qlocker ransomware family has exploited vulnerabilities that allowed full access so that they could deliver ransomware payload to the user's NAS devices [118]. Another attack on NAS devices supplied by QNAP has been done by Deadbolt [119]. The Deadbolt ransomware attack used a zero-day vulnerability that was on the Photo Station Security Update. QNAP is not the only company that has been attacked in the NAS market. The users of Synology, another supplier of NAS devices, had been attacked in 2014 by exploiting a flaw in the unpatched Linux-based DiskStation Manager [111].

Table 3.3 gives the vulnerabilities with the CVSS (Common Vulnerability Scoring System) scores higher than 7 that the popular backup software and hardware providers were exposed to after 2020. The CVE denotation under vulnerability ID means Common Vulnerabilities and Exposures. The table only lists vulnerabilities that had a complete compromise of access and system integrity. It is important to note that all of these vulnerabilities have been fixed shortly after they were recognized. Although any machine that has not been updated will still be open for vulnerability. [132] notes that 76% of vulnerabilities that ransomware exploit in 2022 had been recognized between 2011 and 2019.

Table 3.3: Vulnerabilities Exploited by Ransomware

Vulnerability ID	**Description**	**Date**	**Vendor**	**Citation**
CVE-2021-25270	A local attacker could execute arbitrary code with administrator privileges in HitmanPro.Alert before version Build 901.	08/10/21	Sophos	[121]
CVE-2021-25264	In multiple versions of Sophos Endpoint products for MacOS, a local attacker could execute arbitrary code with administrator privileges.	17/05/21	Sophos	[121]

Table 3.3: Vulnerabilities Exploited by Ransomware

Vulnerability ID	Description	Date	Vendor	Citation
CVE-2020-25223	A remote code execution vulnerability exists in the WebAdmin of Sophos SG UTM	25/09/20	Sophos	[121]
CVE-2021-34996	This vulnerability allows remote attackers to execute arbitrary code on affected installations of Commvault CommCell. Although authentication is required to exploit this vulnerability, the existing authentication mechanism can be bypassed.	13/01/22	Comm-vault	[127]
CVE-2022-26504	Improper authentication in Veeam Backup Replication component allows attackers execute arbitrary code	17/03/22	Veeam	[123]
CVE-2022-26501	Veeam Backup Replication 10.x and 11.x has Incorrect Access Control.	17/03/22	Veeam	[123]
CVE-2020-9478	An OS command injection vulnerability allows an authenticated attacker to remotely execute arbitrary code on Rubrik-managed systems.	13/04/20	Rubrik	[120]
CVE-2021-44057	An improper authentication vulnerability has been reported to affect QNAP device running Photo Station. If exploited, this vulnerability allows attackers to compromise the security of the system.	05/05/22	QNAP	[133]
CVE-2021-44056	An improper authentication vulnerability has been reported to affect QNAP device running Video Station. If exploited, this vulnerability allows attackers to compromise the security of the system.	05/05/22	QNAP	[133]

Table 3.3: Vulnerabilities Exploited by Ransomware

Vulnerability ID	Description	Date	Vendor	Citation
CVE-2020-36198	A command injection vulnerability has been reported to affect certain versions of Malware Remover. If exploited, this vulnerability allows remote attackers to execute arbitrary commands.	13/05/21	QNAP	[133]
CVE-2021-44142	A remote attacker with write access to extended file attributes can execute arbitrary code with the privileges of smbd (Samba file sharing server), and typically root.	21/02/22	Synology	[134]
CVE-2021-29090	Improper neutralization of special elements used in an SQL command ('SQL Injection') vulnerability in PHP component in Synology Photo Station before 6.8.14-3500 allows remote authenticated users to execute arbitrary SQL command via unspecified vectors.	02/06/21	Synology	[134]
CVE-2021-29083	Improper neutralization of special elements used in an OS command in SYNO.Core.Network.PPPoE in Synology DiskStation Manager (DSM) before 6.2.3-25426-3 allows remote authenticated users to execute arbitrary code via realname parameter.	01/04/21	Synology	[134]
CVE-2021-3156	Sudo before 1.9.5p2 contains an error which allows privilege escalation to root via the command ”sudoedit -s” and a command-line argument that ends with a single backslash character.	26/01/21	Synology	[134]

Table 3.3: Vulnerabilities Exploited by Ransomware

Vulnerability ID	Description	Date	Vendor	Citation
CVE-2020-27660	SQL injection vulnerability in request.cgi in Synology SafeAccess allows remote attackers to execute arbitrary SQL commands via the domain parameter.	12/04/22	Synology	[134]

3.2.3. Utilizing Service Utilities in OS

Ransomware attackers use many known service utilities that are legitimate programs, but can be exploited at the wrong hands who have administrator privileges. This is a very common attack pattern and is conducted similarly in most ransomware families. The commands utilized can actually be found in the OS documentation as these are commands that are created for the use of administrators who need to configure the system.

Most known variants belong to the Windows OS. These utilities are; the Volume Shadow Copy Service (Vssadmin), Windows Backup Administrator (wbadmin), Boot Configuration Data Editor (Bcdedit), Windows Recovery Environment (REAgentC), Windows Restart Manager, and Windows Management Instrumentation Command-line (wmic).

The Volume Shadow Copy Service

The Volume Shadow Copy Service allows the creation, listing, and deletion of shadow copies. The vssadmin could also be utilized to manage storage associations which determine where the shadow copies are going to be stored. Furthermore, the tool could be used to change the global settings of the shadow copies, such as the maximum size of the shadow copy storage area, and the maximum number of shadow copies that will be created.

The shadow copy deletion attack is noted under [98, 99, 100, 103, 104, 110, 114, 116]. Research shows that the ransomware families who are reported to use this method are; Avaddon, Babuk, BitPaymer, Black Basta, BlackCat, Clop, Conficker, Conti, DarkWatchman, DEATHRANSOM, Diavol, EKANS, FIVEHANDS, H1N1, HELLOKITTY, HermeticWiper, InvisiMole, JCry, Maze, MegaCortex, Meteor, Netwalker, Olympic Destroyer, Prestige, Prolock, Pysa, Ragnar Locker, REvil, RobbinHood, Royal, Ryuk, WannaCry, and WastedLocker.

This backup destruction method is utilizable after having admin rights on the vssadmin service The command used here is:

```
vssadmin.exe delete shadows /all /quiet
```

[135] notes an interesting new observation in Ransomware attacks. The system will delete the existing shadow copies and will not create new ones, if there is not enough storage allocated for the shadow copies. In order to stop shadow copies from being made attackers resize the maximum amount of storage that is allocated for shadow copy storage. They decrease it to lower or close to 320 megabytes, which is the minimum possible size for a shadow copy. For example the Hakbit ransomware

family [135] and the BlackByte, Conti and Clop ransomware families [98] utilize vssadmin.exe to input the command:

```
”vssadmin resize shadowstorage /for=c: /on=c: /maxsize=401MB”
```

Windows Backup Administrator Utility

The Windows Backup Administrator (wbadmin) is a command-line utility that allows various backup and recovery tasks. In design, it is a utility tool that allows the management of the Windows Server Backup feature. This feature allows the creation and restoration of backups of files, folders, volumes and system states (snapshots). To utilize the wbadmin you need admin or backup operator privileges in the system. According to [136], a simple command that deletes the system backups in the wbadmin service is:

```
”wbadmin delete systemstatebackup -keepVersions:<numberofcopies> | -
version:<versionidentifier> | -deleteoldest [-backupTarget: <volumename>]
[-machine:<backupmachinename>] [-quiet].”
```

[99] notes a similar command that can delete the Windows Backup Catalog:

```
wbadmin.exe delete catalog -quiet
```

Boot Configuration Data Editor

The Boot Configuration Data Editor (bcdedit) on the other hand is a command-line tool that allows users to view, modify and configure settings in the Boot Configuration Data (BCD) store. This store is a database that holds information about the configuration, which notes the boot options, boot menu entries and other boot settings.

The BCD is used to manage multiple OS that are installed on a computer. The administrators of the system then are able to control the booting processes utilizing the tool. The BCD store is backed up and restored through the bcdedit.exe command line. Ransomware groups utilize the bcdedit.exe to disable automatic Windows recovery features by modifying the boot configuration data [99]. They do this by utilizing the command:

```
bcdedit.exe /set default bootstatuspolicy ignoreallfailures & bcdedit /set de-
fault recoveryenabled no
```

Windows Recovery Environment Agent (REAgentC)

The ransomware attacks also target the Windows Recovery Environment Agent [99]. The Windows Recovery Environment is provided by Windows in order to allow users to be able to recover their systems in case of critical issues. The REAgentC tool allows commands to configure, enable, disable, and perform various operations related to the Recovery Environment. The agent can also test the Recovery Environment in order to verify whether it is set up correctly. [137] notes a simple way to disable the Recovery Environment is ”Reagentc /disable”, although there are no explicit examples of the tactic here it could be seen that the ransomware attackers can use ”Reagentc /info” to be able to find where the recovery images are. If they have privileges then they can go and delete them manually. Or they could use the following command:

```
”/setreimage /path <path_to_Windows_RE_image>
[/target <path_to_offline_image>]”
```

This command is able to change where the recovery images are backed up. The backups could be sent to a less safe directory in order for the ease of access of the attackers.

The ransomware Conti [100], Royal, Babuk, and Lockbit [116] use the Windows Restart Manager to close applications and services that are running. This allows for two things. First, the data that is handled by the applications are free to be encrypted. Second, if there are any backup or anti-ransomware software, they are also terminated. The Windows Restart Manager is noted [116] to be utilized maliciously in the following way; the session is started through the API (Application Programming Interface) Call RmStartSession, then the API Call RMRegisterResources registers the targeted files, RMGetList shows which applications or services are handling the resource, finally, RmShutDown is called in order to terminate the applications and services which use the resource.

Windows Management Instrumentation

The Windows Management Instrumentation Command Line (wmic) is a tool that allows users to interact with the Windows Management Instrumentation infrastructure. This is a management technology that provides a standardized way to perform administrative tasks and to manage resources in a Windows OS. The tool is quite powerful, and if the ransomware group gets a hold of it, it is quite bad news. The tool can access and manage critical infrastructure such as; system information, software management, hardware management, process management, and user account management.

The system information relates to the computer name, OS version, processor details, memory configuration and etc. This information can highlight to an attacker which vulnerabilities may be present in the corresponding version of the system. The software management allows the attacker to be able to see installed applications, and their versions. The attacker can also then uninstall certain applications.The hardware management on the other hand includes information about the devices, disk drives, network adapters, printers, and other hardware. This is basically the functionality of the ”device manager” program that is found in the Windows OS. The attacker also can retrieve information about running processes and the details about the process properties. They can also start or terminate processes as they see fit. With the user account management functionality, they can manage local user accounts on a system. Here they can create new user accounts, or modify certain account properties.

Ransomware families such as BlackCat, BlackByte, Maze, Trickbot, BlackMatter, LockBit, and Nefilim utilize the command ”wmic shadowcopy delete /nointeractive” to delete the shadow copies [99, 107, 135].

Ransomware such as the Darkside, and Netwalker families also utilize the PowerShell to delete the shadow copies [135];

```
”Get-WmiObject Win32_Shadowcopy | ForEach-Object $_.Delete();”
```

3.2.4. Identifying and Terminating Processes

Ransomware utilizing these command-based attacks also closes down known antivirus, security, database, and backup software instances. These processes and service lists get updated and may change from ransomware to ransomware. They might also change their commands with respect to the network that they have scoped out during the reconnaissance stage. [98] notes that they target things that contain the names ”vss”, ”sql”, ”oracle”, ”veeam”, and ”backup”.

They utilize the following arguments in their commands to stop services and processes [98]:

- “net stop “<service_name” /y>”
- “sc config “<service_name” start= disabled” • “iisreset.exe /stop”
- “taskkill” and others

The list of processes terminated, compiled from [98, 100, 102, 103, 104, 105] are listed below;

Table 3.4: List of Services and Processes Terminated

Names of Services and Processes Terminated	
AcrSch2Svc	Intuit.QuickBooks.FCS
AcronisAgent	memtas
BackupExecAgentAccelerator	mepocs
BackupExecAgentBrowser	PDVFSService
BackupExecDiveciMediaService	QBFCService
BackupExecJobEngine	QBIDPService
BackupExecManagementService	QBCFMonitorService
BackupExecRPCService	RTVscan
BackupExecVSSProvider	SavRoam
CAARCUpdateSvc	sophos
CASAD2DWebSvc	sql
ccEvtMgr	stc_raw_agent
ccSetMgr	svc
DefWatch	VeeamDeploymentService
GxBlr	VeeamNFSSvc
GxClMgr	VeeamTransportSvc
GxCVD	veeam
GxFWD	VSNAPVSS
GxVss	YooBackup
backup	YooIT
vss	zhudongfangyu

3.2.5. Backup Integrity & Attacks to Storage Management

Following the prior sections, it could be seen that ransomware has multiple ways to attack backups. This section is more related to what they do with the backups once they have access. They mainly can corrupt, delete, encrypt, or infect backups. Corruption makes it so that the backup files are inoperable during recovery. The corruption angle of backups has been noted in [96] and [97]. Some ransomware attacks the management system of the backups rather than the files themselves. They change them in order to create garbage and unusable backups which is quite related to corruption. This is called backup poisoning [96].

The deletion and encryption of the backup files for both the data and the system backups are mentioned in [96, 98, 99, 109, 138].

Infection on the other hand is more sinister. [96] and [139] note that if an environment is rebuilt while battling an infection, the malware might be restored with the environment. This way, the attacker can quickly seize control again. This is also noted as another form of poisoning [96].

3.3. Backup & Recovery Practices

This section focuses on the backup & recovery practices that have been found in 102 documents. 73 of the documents have had explicit mentions of either a backup or a recovery practice. The important thing to note here is that not all sources are created equal. Experience in the field indicates that most of the knowledge in the industry is predated by established and credible cybersecurity frameworks. This section will first report the built codebook, the quantification of the codes, their analysis, and the abstraction. Then the narrative will delve into each category in the codebook.

11 popular cybersecurity frameworks have been utilized in order to generate the basis of the codebook to be expanded further during the analysis of additional sources. The expanded codebook is thus the summary of all recommendations and answers the second sub-question. The cybersecurity frameworks used in the generation of the codebook inductively are;

- National Institute of Standards and Technology (NIST) Cybersecurity Framework
- International Organization for Standardization (ISO) 27001 & 27002
- Center for Internet Security (CIS) Controls
- Digital Operational Resilience Act (DORA)
- Information Systems Audit and Control Association (ISACA) COBIT 4.1
- German Federal Office for Information Security IT-Grundschutz
- North American Electric Reliability Corporation Critical Infrastructure Protection (NERC CIP) Cybersecurity Standards
- Cybersecurity Maturity Model Certification (CMMC)
- Australian Government Information Security Manual (ISM)
- European Union Agency For Cybersecurity (ENISA) Joint Publication
- Health Insurance Portability and Accountability Act (HIPAA) Cybersecurity Guidance Material

A short explanation on the background and the relevance of the frameworks is then necessary. They are given in the following paragraphs in the order that they have been established in the list above.

3.3.1. Introduction to the Cybersecurity Frameworks

National Institute of Standards and Technology (NIST) Cybersecurity Framework

The NIST framework is created by the National Institute of Standards and Technology, part of the US Department of Commerce. It was established after the President issued Executive Order (EO) 13636, named *"Improving Critical Infrastructure Cybersecurity"* in February 2013 [140], to ensure the reliable function of the critical infrastructure in the US so that national and economic security is protected [141]. Multiple stakeholders, including the NIST have generated the framework to establish the best standards, guidelines, and practices in order to reduce cyber risks. There is however one more publication that is of importance also from NIST, named the NIST Special Publication (SP) 800, where articles 4, 6, and 9 are specifically related to data backups and are referenced in the main NIST Cybersecurity Framework.

ISO/IEC 27001 & 27002

The ISO/IEC 27000 family of standards that are known for giving guidelines to organizations regarding information security management. ISO stands for the International Organization for Standardization and IEC, International Electrotechnical Commission. Their standards are one of the most essential works to establish information security in organizations. The latest version of ISO/IEC 27000 has come out in 2018. The family 27001 is specifically related to cybersecurity and delves into data backup procedures against a cyber attack. It is marketed as a tool for risk management, better cyber-resilience and achieving operational excellence [91].

Center for Internet Security (CIS) Critical Security Controls (CSC)

The Center for Internet Security establishes critical security controls in order to allow organizations to have access to best practices they can utilize to strengthen their cybersecurity infrastructures [142]. The CSC version 8 Denotes 18 controls and CIS Safeguards. They categorize different implementation groups, from small-scale organizations to high risk financial enterprises. They identify Asset Types and Security Functions coupled with each safeguard. They denote the backups in control 11 named, Data Recovery.

Digital Operational Resilience Act (DORA)

The European Parliament and the Council of the European Union, have created and adopted the new Digital Operational Resilience Act (DORA) regulations [92]. The aim of the regulations are to make sure that the financial institutions, who with their downfall can jeopardize the financial system, are properly secured against growing threats. DORA provides rules for the "protection, detection, containment, recovery, and repair capabilities against ICT-related (incl. cyberattack) incidents" [143]. Thus, organizations must be vigilant toward their recovery capabilities. The DORA regulations are expected to enter into force in 2025.

Information Systems Audit and Control Association (ISACA) COBIT 4.1

The Information Systems Audit and Control Association (ISACA) is a non-profit entity that develops Information Systems (IS) auditing and control standards. They name themselves as a leading global provider of knowledge. They offer credible certifications, advocacy, and education on IS security, and enterprise IT governance. COBIT

4.1 is their work that aims to educate professionals on the governance of enterprise IT (GEIT).

COBIT 4.1 is a comprehensive framework that assists organizations in ensuring that the governance and management of organization IT objectives are met. They offer a holistic view and the framework is targeted towards all sizes of organizations. They define maturity levels with regard to different levels of adherence to rules. There are 5 maturity levels namely (with their corresponding ranks); non-existent, ad hoc, repeatable but intuitive, defined, managed and measurable, and optimized.

German Federal Office for Information Security IT-Grundshutz

The IT-Grundschutz [90] by the German Federal Office for Information Security (BSI) is a risk management framework that is aimed to serve as a benchmark for all organizations from all levels. It includes guidelines for data backup and recovery to protect critical information assets. It has a total of 104 modules. The IT Grundshutz notes the backup concept under the section ”CON.3. Data backups” the backups are established to be essential in contingency planning and business continuity of organizations.

North American Electric Reliability Corporation Critical Infrastructure Protection (NERC CIP) Cybersecurity Standards

North American Electric Reliability Corporation Critical Infrastructure Protection (NERC CIP) Cybersecurity standards [144] has been created in order to regulate, monitor and manage the security, specifically cybersecurity, of the Bulk Electric System (BES) in North America [145]. The plan is a regulation and thus is a mandatory requirement for organizations that fall under the BES umbrella.

Cybersecurity Maturity Model Certification (CMMC)

The Cybersecurity Maturity Model Certification (CMMC) [146] is a set of requirements created by the US Department of Defense in 2021. The framework is aligned with other cybersecurity frameworks such as the NIST 800 SP, in order to establish strong cybersecurity standards for adoption to protect critical assets and sensitive information. CMMC 2.0 framework consists of 3 maturity levels and 14 cyber security domains. Among these domains, the Media Protection domain specifically relates to the protection of data backups.

Australian Government Information Security Manual (ISM)

The Information Security Manual (ISM) [147] is created by the Australian Government's sub-entity, the Australian Cyber Security Centre. It establishes recommendations and guidelines for system management. For data backups, they note 5 main recommendations and establish their corresponding controls.

European Union Agency For Cybersecurity (ENISA) Joint Publication

The European Union Agency for Cybersecurity (ENISA) has partnered up with CERT-EU to establish cybersecurity best practices so that all organizations, both in the public and private sector can apply in order to increase their cybersecurity posture and resilience against attacks. In their joint publication, article 6 relates to data backup strategies. Other articles on the other hand relate more to the protection of the backups.

Health Insurance Portability and Accountability Act (HIPAA) Cybersecurity Guidance Material

The US Department of Health and Human Services has created the cybersecurity guidance framework in order to assist organizations, focusing on the healthcare sector, who are covered under the Health Insurance Portability and Accountability Act (HIPAA). The cybersecurity material is explicitly related to stopping ransomware and aims to make organizations more resilient by teaching them the best practices to prevent and recover from ransomware attacks.

3.3.2. Cybersecurity Framework Best Practices

The cybersecurity practices that have been recommended explicitly regarding backup strategies and recovery per framework have been given below in the Table 3.5.

Table 3.5: Best Practices from Frameworks

Framework	**Backup Strategy**
NIST Framework	• Achieve Network Segmentation • Establish Data Plane Seperation (incl. physical off-site) • Create documentation and monitor data backup policies • Utilize the 3-2-1 Backup Strategy • Have an alternate storage site • Have different criticality levels for different data and document their corresponding backup plan • Save more than one copy of backups • Have frequent and accessible backups • The backups must be tested, and if recovery time is important, also be tested end-to-end to track time until full recovery. • Monitor backup file lifecycle through a recovery catalogue, including malware analysis to prevent ransomware poisoning • Establish logical and physical controls over the data backups • Set up proper access management and authorization
ISO/IEC 27001 \ 27002	• Produce accurate and complete records of the lifecycle of the backup copies and their corresponding restoration processes. • Establish criticality of data and systems. • Establish recovery point, time objectives. • Store the backups in an off-site location with physical and logical controls. • Backup regularly. • Test the backups regularly to see if they are fit for use. • Encrypt the backups. • Monitor backups to prevent ransomware or malware infiltration or corruption.

Continued on next page

Table 3.5: Best Practices from Frameworks (Continued)

CIS Controls v8	• Report and Document Backup Plans and Data Recovery Processes. • Ensure Regular Automated Backups. • Test Backups Regularly. • Protect Backups through physical security or encryption in any media form. • Ensure that there is at least one offline (off-site) backup destination. • Generate complete system Backups, as well as critical (sensitive) data backups.
DORA	• Ensuring regular and frequent backups. • Ensuring regular testing of backups. • Indicate the scope of the backups, and establish criticality metrics for different data and systems. • Ensure that there is physical or logical segregation of backups. • Ensure proper access management to the backup storage. • Create redundant backups.
ISACA COBIT 4.1	• Produce backup and recovery procedures. • Establish criticality of data and systems, and their dependencies. • Store the backups in an off-site location. • Backup regularly. • Monitor backups. • Test the backups regularly to see if they are fit for use. • Encrypt the backups.
IT-Grundschutz	• Utilize data protection frameworks. • Establish data backup procedures. • Document and report backup plans. • Backup regularly and frequently. • Ensure correct access management, authorization for storage media. • Ensure overwriting controls on the backups. • Monitor and test backups frequently. • Encrypt backups.
NERC CIP	• Document systems and data backups, and their corresponding storage media. • Test backups (and recovery) regularly.

Continued on next page

Table 3.5: Best Practices from Frameworks (Continued)

CMMC	• Regular, comprehensive, and resilient backups are essential. • Resilience of the data backups can be achieved by using off-site and offline storage locations. • Data backups must be ensured to be made available during recovery. • To ensure redundancy, more data backups must be created. • Test backups (and recovery) regularly.
ACSC ISM	• Establish data backup and recovery process and procedures. • Perform backups frequently and retain according to organizational goals. • Ensure proper access management and authorization for the backup data. • Ensure that backup data cannot be modified nor deleted without proper access during their retention period. • Backups must be regularly tested for recovery.
ENISA Joint Publication	• Set up Multi-Factor Authentication • Ensure proper access management to backups through controls and logging. • Control access to internal networks and systems • Employ network segmentation Block or severely limit internet access for servers that have the backups • Change all default credentials • Ensure proper patch management and that software is up-to-date • Harden the safety controls of cloud environments that contain the backup data • Review organizational data backup strategy • Align data backup strategy with setting RTOs and RPOs that fit business needs Frequently backup data, especially the critical data
HIPAA	• Perform backups frequently. • Verify the integrity of backup data through testing. • Do end-to-end testing to increase confidence in full recovery. • Maintain backups offsite and offline. • Plan, maintain, implement, and test a data backup plan. • Analyze Criticality of Applications and Data.

The table gives quite an overview on what the state-of-the-art, best practices regarding backup & recovery are. These directly relate to recommendations to be made for organizations. This overview establishes the basis of recommendations made in Chapter 5 for the challenges and the ransomware attacks that have been discovered from all of the methods in the thesis. These best practices are expanded on the following sections.

3.3.3. Best Practices Expanded from Reports & Other Material

Additional sources have built up upon the initial codebook introduced with the cybersecurity frameworks to make it more comprehensive and generalized. The expanded codebook that has been created through analyis, for backup & recovery practices can be found in Table 3.6 and Table 3.7. The expanded codebook answers the second subquestion. The ”#” column on the codebook represents the number of times a particular concept was mentioned and is cumulative for top-level categories.

Table 3.6: Codebook of Operational Practices & Procedures in Backups and Recovery

Category/Subcategory		#	Citations
Generation		63	[90, 92, 93, 94, 95, 96, 97, 98, 99, 100, 105, 106, 107, 109, 112, 113, 114, 119, 125, 138, 139, 142, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183]
	Isolate Recovery Data	47	[90, 92, 94, 95, 96, 97, 98, 99, 105, 106, 107, 109, 112, 113, 114, 125, 138, 139, 142, 146, 147, 148, 149, 150, 151, 152, 156, 157, 159, 160, 162, 163, 164, 166, 167, 168, 169, 170, 171, 174, 176, 179, 180, 181, 184, 185]

Table 3.6: Codebook of Operational Practices & Procedures in Backups and Recovery

Category/Subcategory/Properties			#	Citations
		Use Offline, Cloud or Offsite Systems and Services	37	[92, 94, 95, 96, 97, 98, 99, 105, 106, 107, 112, 113, 125, 138, 139, 142, 146, 149, 150, 152, 157, 164, 166, 167, 168, 169, 170, 171, 176, 177, 179, 180, 181, 185]
		Airgapping	6	[106, 156, 159, 160, 162, 168]
		Manage and Periodically Assess Offsite Arrangements	4	[90, 96, 147, 152]
	Regular Backups		28	[90, 93, 100, 107, 112, 113, 119, 138, 139, 142, 146, 147, 149, 152, 155, 159, 163, 166, 168, 171, 172, 176, 177, 178, 180, 182]
	3-2-1 Strategy		17	[98, 107, 139, 153, 154, 160, 161, 167, 168, 176, 178, 181, 182, 183, 186]
	Immutable Backups		14	[90, 94, 96, 97, 113, 147, 151, 162, 166, 167, 168, 169, 173]
	Ensure Compatability Correct Configuration		3	[147, 152, 175]
	Establish Failover Backup Systems		3	[92, 97, 166]
	Redundant Backups		3	[109, 114, 170]
	4-3-2-1 Strategy		1	[165]
	Continuous Data Protection (CDP) Backups		1	[96]
	Ensure at least one Full Backup		1	[96]

Table 3.6: Codebook of Operational Practices & Procedures in Backups and Recovery

Category/Subcategory/Properties			#	Citations
Protection			52	[90, 93, 94, 95, 96, 98, 100, 104, 105, 107, 108, 109, 112, 113, 114, 117, 118, 119, 125, 138, 142, 146, 147, 148, 149, 151, 152, 155, 157, 159, 161, 162, 163, 164, 166, 169, 170, 171, 172, 173, 175, 176, 177, 180, 182, 184]
	Access and Privilege management		28	[90, 93, 94, 96, 98, 107, 108, 109, 113, 114, 117, 125, 146, 148, 151, 152, 155, 159, 162, 169, 173, 177, 184]
		Implement MFA	13	[93, 94, 98, 107, 108, 109, 114, 125, 148, 162, 169, 177]
		Minimum Access Rights and Privileges	11	[90, 96, 98, 109, 113, 114, 117, 159, 162, 169, 177]
		Zero Trust Access Strategy	5	[148, 151, 159, 162, 173]
		Create Quorums	2	[151, 184]
		Physically Secure Devices and Media	2	[146, 152]
		Seperation of Duty	2	[96, 184]
		Disable Plaintext Passwords	1	[109]
		Enable Credential Guard	1	[114]
		Encrypt Login Credentials at Rest	1	[96]
		Instead of passwords use SSH	1	[125]
		Proper Authentication	1	[155]
		Role-based Access Controls	1	[94]

Table 3.6: Codebook of Operational Practices & Procedures in Backups and Recovery

Category/Subcategory/Properties			#	Citations
	Encrypting Backups		18	[90, 94, 98, 107, 113, 142, 146, 149, 155, 162, 163, 166, 169, 171, 173, 176, 184]
		Key Management	5	[90, 146, 149, 155, 169]
	Monitor Domains, Backups Recovery Data and Application Health		17	[92, 93, 94, 96, 104, 107, 113, 114, 117, 151, 157, 160, 163, 170, 175, 179, 184]
	Network and Data Plane Segmentation and Seperation		17	[95, 96, 98, 100, 105, 107, 108, 112, 113, 125, 142, 155, 159, 161, 169, 170, 177]
	Patch Management		17	[98, 100, 105, 107, 109, 113, 117, 118, 119, 130, 153, 159, 164, 169, 170, 172, 177]
	Disable all Unnecessary Services and Protocols		5	[96, 99, 107, 109, 177]
	Streamline IT Landscape		4	[94, 151, 178, 180]
	Utilize Soft Delete		1	[173]
Testing Recovery Employee Training			28	[90, 92, 96, 98, 107, 108, 138, 142, 147, 148, 149, 152, 153, 154, 155, 159, 160, 162, 163, 166, 171, 172, 176, 179, 180, 182, 184]

Table 3.6: Codebook of Operational Practices & Procedures in Backups and Recovery

Category/Subcategory/Properties		#	Citations
Backup and Recovery Policy		27	[90, 92, 96, 98, 108, 109, 113, 114, 117, 138, 142, 144, 147, 149, 151, 152, 153, 155, 159, 163, 172, 176, 178, 179, 183, 187, 188]
	Content, Criticality, Prioritization and Characteristics of Backups	9	[92, 96, 117, 142, 149, 152, 178, 179]
	Establish Operational Plans Regarding Saving Data	5	[113, 114, 153, 172, 183]
	Periodic Audits	5	[96, 109, 113, 114, 151]
	Establish Responsibility	4	[149, 178, 187, 188]
	Keep inventory of Backup Storage	4	[90, 96, 142, 152]
	Establish Communication and Cooperation Plan	2	[98, 163]
	Establish KPIs and Requirements	1	[179]
	Lifecycle Management	1	[179]
	Tiered Approach	1	[178]
Utilizing SaaS and BaaS		8	[94, 168, 169, 173, 180, 181, 188, 189]
Data Reduction		3	[96, 179, 186]
	Data Deduplication	3	[96, 179, 186]
	Utilize Data Compression	1	[96]
Recovery Principles		2	[142, 147]
	Use a Version of Backup that Predates Infection	1	[142]
	Wipe and Rebuild From Good Backup	1	[147]
Collaboration		1	[151]

The table will be analyzed by focusing on the largely focused concepts to understand what is the main trends and mentions in the material. In the operational practices & procedures table, it could be seen that in the material most focus is given to backup generation (63 mentions), followed by protection (50 mentions), testing and training (28 mentions), and establishing backup and recovery policy (27 mentions). These are the main discussion points in the material with regard to the operational practices. Further mentioned categories have a large gap with respect to the number of mentions and are minuscule comparatively. These are utilizing Software as a Service (SaaS) and Backup as a Service (BaaS) utilities (8 mentions), utilizing data reduction methods

(3 mentions), administrating recovery principles (2 mentions), and collaboration (1 mention).

Under backup generation, the most amount of mentions are made with respect to isolating the backup data through storing them in an offsite, cloud, air-gapped, or offline environment (47 mentions). This is followed by the mention of the necessity of (frequent) regular backups (28 mentions). Then comes utilizing the 3-2-1 strategy (17 mentions) and immutable backups (14 mentions) the other categories are quite minuscule in the number of mentions made.

Protection on the other hand near equally mentions: encryption of backup data (18 mentions), monitoring (17 mentions), network and data segmentation and separation (17 mentions), patch management (17 mentions), and access and privilege management (15 mentions). The others are comparatively lower in the number of mentions.

The concepts under the Backup and recovery policy category are primarily about the need for documentation and the characteristics that these policies must have. The most mentions in the subcategories of the backup and recovery policy category are made on; deciding on the content, data criticality, recovery prioritization, and the characteristics of backups (9 mentions). This is followed by establishing operational plans with regard to saving backup data (5 mentions), ensuring the establishment and documentation of periodic audits (5 mentions), establishing responsibility and accountability (4 mentions), and keeping an inventory of backup storage (4 mentions).

Table 3.7: Codebook of non-cloud Backup Storage Options in Backups and Recovery

Category/Subcategory/Properties			#	Citations
Non-cloud Backup Storage Options			12	[90, 139, 146, 150, 162, 169, 174, 176, 179, 182, 190]
	Mobile Removable Media		5	[90, 146, 174, 179, 190]
		Tape Media	5	[90, 146, 174, 179, 190]
		Optical Disk	2	[90, 146]
		External Hard Drives	1	[146]
		USB Drives	1	[146]
	Networked Attached Storage		2	[119, 146]

The non-cloud backup storage options table is relatively balanced and give mentions to mobile removable media (5 mentions) and its subcategories; Tape media (5 mentions), Optical Disk (2 mentions), External Hard Drive (1 mentions), and USB Drives(1 mentions). Followed by Networked Attached Storage (2 mentions). Albeit most material notes cloud as a primary offsite arrangement, 12 material thus specifically highlight the usage of traditional storage media for keeping backups. The abstraction that has been formulated for the backup & recovery practices are given in Figure 3.2. It identifies key categories that is observed in the material.

The following sections then will be the manifestation of the analysis of each category that has been noted on the abstraction figure. The backup & recovery practices

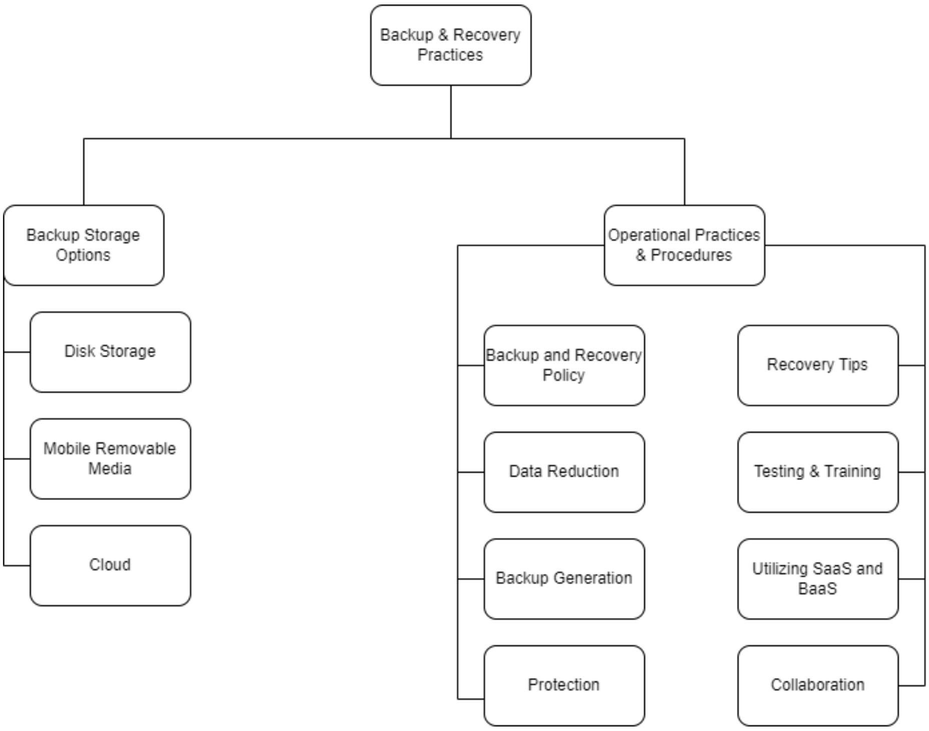


Figure 3.2: Abstraction of Backup & Recovery Practices

are grouped under main two categories; namely the backup storage options, and operational practices & procedures. Backup Storage options contain choices on disk storage, mobile removable media, and the cloud. On the other hand, the operational practices and procedures delve into the subcategories; backup and recovery policy, data reduction, backup generation, data protection, testing & training, utilizing software as a service (SaaS) and backup as a service solutions (BaaS), finalizing with recovery tips and collaboration. It is important to note that not all citations will be listed in their respective categories as having 70 citations noting that regular backups must be made is not beneficial for further exploration. For information regarding how many of the material has referenced a specific strategy please refer back to Table 3.6 and Table 3.7.

3.3.4. Backup Storage Options

Organizations can select from a plethora of backup storage options in order to store their backups. All these might suit the organizational needs with combined with what other measures are utilized. The main backup storage options are; Hard disk drives (HDDs), Solid-state drives (SSDs), Flash Drives, Optical Disks (CD-ROM, DVD-ROM, and Blu-rays), Network Attached Storage (NAS), Storage Area Networks (SAN), Tapes, and Cloud Storages. This section is subdivided into Mobile Removable Media, NAS, Disk Storage, and the Cloud. Even though the citations regarding the cloud have been coded under offsite isolation strategy, the cloud as a storage media will be mentioned here for completeness. The Mobile Removable media contains tapes, optical disks, external hard drives, and USB drives.

Here, storage tiering is a notable mention [191]. The storage tiering allows for the storing of data with regard to their purpose and how often they are accessed.

Mobile Removable Media

Magnetic tapes have been used as a cost-effective storage media since a long time ago. Tape backups do offer some major benefits: they are cost-effective due to the widely adopted non-proprietary Linear Tape Open (LTO) standard, they're easy to physically transport, and they are compatible with many backup applications due to their long history as a backup medium [179].

The ease of transport allows organizations to quickly establish an offsite and offline measure for their backups providing security against ransomware attacks especially if they are helped by a physical backup security provider such as Iron Mountain [190]. [174] notes that the tapes offer a true physical gap between the primary domain and the backup files. The tapes do need environmental and physical protection which is much easier to monitor and keep secure from ransomware.

Both [146] and [90] note Optical Disks as a way to store backups as they could be made immutable, and read-only. The benefits are that they are cheap, easy to find and use and the storage is quite simple. However, they have a few drawbacks. Many providers are pulling away from optical disks and newer applications are made with no optical disk drives [192]. Furthermore, the data size in organizations may outscale the capabilities of the disks. Quadruple-layer Blu-ray Disks can allow for a maximum of 128 GB of data.

Network Attached Storage

Organizations could utilize NAS systems for their backups. The NAS devices are high-capacity storages that can connect to networks so that specified users can access the files without having the need to plug them into another device like a hard disk. The NAS devices come with their respective properties relating to their speed and compatibilities, etc. Two big providers for the NAS are QNAP and Synology. [119] notes that QNAP had stated that people should not connect their NAS directly to the internet and instead use VPN or the myQNAPcloud Link so that ransomware attackers cannot access the machines. The NAS devices must be kept behind a firewall with strong anti-intrusion settings and never be exposed to the internet [119].

Cloud Storage

The cloud storage is noted especially with respect to it offering cost benefits to organizations who would like to keep their data offsite. Therefore this is not the only section that the solution will be discussed in. The Cloud is basically connecting to a server that is hosted by a cloud provider who handles the storage infrastructure instead of an organization themselves. The cloud solutions could also be personalized, and some offer integrated backup and recovery options. Some also provide interesting data vault concepts where backups are immutable, read-only, or have other concepts such as multi-factor authentication etc. that make backups harder to access to attackers. The primary business model in cloud backups is pay-as-you-go and the best-known providers are Microsoft, Amazon, and Google.

3.3.5. Operational Practices & Procedures

Under this section, operational practices & procedures such as; establishing backup and recovery policies, backup generation, protection, data reduction methods, recovery principles, testing & training, utilizing BaaS solutions, and collaboration will be examined further.

Backup and Recovery Plans

Backup and recovery plans are critical to establish so that organizations know what, where, how and when during a cyber attack scenario. These plans must be well documented, and exercised so that all employees are aware of their responsibilities and objectives to restore the infrastructure back. The backup and recovery plans must include certain details.

[187] explains that the requirements, processes, policies, and procedures for operating the backup environment, and the corresponding roles, responsibilities and accountability must be documented in the backup and recovery plans. There could also be notation of a recovery team who have specific roles and responsibilities during a crisis [149]. [188] notes that many organizations do not back up their data as they think the cloud providers have that covered. However, cloud providers mainly work on a shared responsibility model where the backup and recovery responsibilities fall to the organization. Thus, these must be well documented, and organizations must establish in their own backup and recovery plans where specifically their responsibilities begin.

Furthermore, another important point is that organizations must ensure they establish criticality metrics in order to map out their critical data and applications [149]. CIS controls v8 [142] note that the backup procedures need to be aligned with the data value, sensitivity, and retention requirements. This is so that recovery prioritization could be done during recovery based on such documentation. The prioritization should also document what is the core business functions that are planned to be recovered first. The documentation then must also ensure to include the dependency between different databases and data, applications, etc. Following dependency, the sequentiality or write-order fidelity must also be protected and specific consideration must be given during the documentation process, especially in organizations where transactions have increased risks. Additionally, [117] states that every backup & recovery plans must ensure to have an inventory of the assets and data belonging to the organization, including the authorized and unauthorized devices and software. This will allow both the recovery team and the analysis team to have a strong oversight during a crisis and reduce downtime.

Moreover, organizations must elaborate on the technical requirements of the backups. [90] & [187] note that, organizations must check their backup windows, recovery point objectives(RPOs), recovery time objectives (RTOs), retention requirements with respect to their situation. Additionally, [92] notes that the scope of the data that is subject to the backups, their backup frequency, and tiering considerations must also be documented. The tiering in this context means that different backup and recovery policies apply to different categorizations of data. Tiering could be done by establishing different frequencies and retention to data [96]. The backup & recovery procedures must also define the type of backup, the backup schedule, how the storage media is to be used, how the data must be transported, how it will be encrypted and how the key management will be undertaken, and how the backup protection will further be established [90, 96]. Moreover, the content of the backup storage must also be determined and documented in collaboration between business process owners and the IT personnel [152]. These will then allow organizations to assess whether their objectives and the solutions they utilize align. The documented plans then also allow for organizations to be able to consider a mix of products and service providers in order to establish an effective backup & recovery system.

[90] states other unique considerations that must be added to the backup and recovery plans. These are to document the; storage spaces of the storage media that has been selected, the changing times, availability requirements, confidentiality requirements, integrity requirements, legal requirements, and requirements for deletion of backups. [90] further notes that, the documentation of the backup & recovery plans should include access management considerations which are; ”the user IDs, user groups, and rights profiles that have been approved and created.” This way, monitoring could be done effectively in order to understand which users have authorization for the backups and whether they are appropriate [90].

The considerations above will ensure that organizations are on the right path with regard to creating a solid backup & recovery plan against ransomware attacks. However, the last step of the puzzle is to ensure the maintenance of the plans. To this end periodical audits (both in-house and perhaps external assessments), reviews, and updates must be done [96, 142].

Backup Generation

In the material, there are established strategies regarding backup generation. Mostly, the focus is on isolating the backup & recovery data through either airgaps, offline, cloud, or offsite; or sustaining regular and immutable backups.

In many ransomware attacks the attackers destroy the backups that are connected to the network [109]. Then what is important here is to make sure that the backup data that is crucial for recovery should not be in the primary domain and must be stored offsite or offline [96, 112]. The NIST Framework [96] describes the offsite concept as ”alternate storage sites are geographically distinct from primary storage sites.” Offsite backup must not be deletable from the main network [105]. The offsite measure extends to utilizing cloud infrastructure as well. If utilized, the backup data sent to the third party will not be part of the primary network of the organization [113]. Utilizing the cloud for ransomware recovery offers reduced costs, increased security, and immediate availability during recovery [174].

The offline, and the isolation parts relate to off-the-network isolation and a use case is the airgaps. Air-gapping is one of the most robust backup mechanisms as it is disconnected from the network and the internet [161]. Although air-gapping primarily relates to a real physical disconnection, there are many third parties who offer specialized solutions in order to create virtual airgaps. The Microsoft Azure backup notes that the backup data will be secured through isolation methods under the Recovery Services or the Backup Vault [184]. A true airgap however is noted to be things that can be physically separated out and locked out from the main network. Examples here are External Hard Drives, Tape solutions, or Disk solutions [106]. The offsite and offline (isolated) backups ensure that ransomware has a hard time laterally moving throughout the network and accessing backup files. Even in the case of access, many cloud and other backup solution providers extend their services to include immutable and redundant copies that only they can access. Then, the provider simply can rollback to a prior version that is not tainted by ransomware.

Another important note is that backups must be done regularly [90], frequently [152], and possibly with respect to a schedule [178]. Automation here could be beneficial for organizations to keep track of and verify the backup data after they are created [180]. [149] states that the principle for regular data backups is daily backups. These regular backups could be made to a media or a remote location such as the cloud. The principle also recommends that the backups must be stored in different media for end-of-week and end-of-month backups. Here, it is seen that there is an implication of a tiering structure of backups where some backups comply with different policies. On this end, additionally [176] notes that not all types of data should always be backed up at the same time. Some critical information however should be necessary to be continuously backed up in real-time, such as healthcare data. This real-time backup could be delivered by the ”Continuous Data Protection (CDP) which is a form of backup that supports a fine-grained recovery and improved recovery point objectives” [96]. Lastly regarding the regular backups, [96] notes that even if an incremental backup type is selected in organizations, regular full backups must be made as the incremental backups are not usable during recovery without their baseline copy.

The organizations must also keep their backups in an immutable state [94]. The immutability refers to the concept that backup data cannot be intentionally or unintentionally altered by any party [90]. The immutability can be established through retention locking, vault locking [96] or using media storage that allows for immutability rules such as write-once-read-many (WORM) [94].The retention locking is the concept where the backup data is immutable until a specific period is up. Vault locking could be realized by a solution offered by IBM. The data vaults allow the organizations to input their sensitive material and once the lock switch is on, the data cannot be further altered. The WORM media include CD-Rs (R stands for recordable), DVD-Rs, and Blu-ray Disks (BD-Rs) [96]. These WORM media is also contrasted to switchable write-protection media by the NIST Framework [96] as the switchable write-protection is not necessarily a write-once media. Write protection is a feature, but WORM is a specific type of storage media. The write-protected media are tapes, CD-RWs (RW stands for re-writable), DVD-RWs, and USB drives. The immutable backups then will make sure that ransomware cannot encrypt, infect or corrupt the backup data that are being retained.

[170] notes that the backup copies must also be made redundant in order to ensure that there is no single point of failure. The redundancy also applies to systems. In critical systems such as financial organizations, etc. failover backup systems must be established [92]. These failover backup systems allow for a switchover from the primary ICT infrastructure to the backup during a crisis. These failovers are established to ensure operational continuty and to mitigate disruptions to operations [166].

Moreover, backups must be made compatible between the hardware and software which are dedicated to the restoration processes [152]. The ACSC Information Security Manual[147] notes that backup data could be made into a pre-documented format that is preferably an open standard to maximize compatibility. The backups must also be configured correctly in the backup solutions as they could fail. [175] notes for example, if the Azure Backup agent is not set up correctly, there will be problems with the backup files being written correctly.

Organizations must also utilize the 3-2-1 strategy which is a conglomeration of several concepts in securing backups. The ”3” means that there must be three copies of data, here one being the production data and two backup copies. The ”2” relates to securing the backup copies on two different media such as disk and tape. The ”1” means that at least one of the copies must be sent offsite. This strategy is noted to mitigate ransomware risk on backups as much as possible [107, 154]. [186] notes that utilizing the 3-2-1 strategy could be made easier by utilizing Disaster Recovery as a Service (DRaaS) platforms as they satisfy all the requirements instead of the organization.

Although not ground-breaking, a unique contender to the 3-2-1 backup strategy is found in [165]. The material notes that the 4-3-2-1 strategy will be more effective in circumventing ransomware attacks against backups. Here [165] notes that there are 4 copies of data, one being the production and three backups. These backup copies are stored in three different sites, one in the cloud, while the other two are up to the organization for local and offline solutions. Then two types of storage are used, one offsite and one on-premise. Lastly, one of the backups is made an immutable copy. Here it could be seen that the main difference is having one more redundant backup copy, using one more storage media, utilizing both on-premise and offsite and lastly an immutable backup copy exists. This method does indeed increase the cyber resiliency against ransomware attacks but does come with increased costs.

Protection

Backup data that is generated for the purposes of recovery must also be kept safe through correct protection mechanisms. The mentioned categories under protection are access and privilege management, encryption, network practices, monitoring, patch management, disabling of services and protocols, and Streamlining the IT Landscape. Here. the most mentions are made with regards to the first five listings, where access and privilege management holds the majority of mentions.

Organizations can adopt the best practice of multifactor authentication [148] for access management. It is a measure that adds an extra layer of protection during the authentication process. Multifactor authentication usually utilizes a combination of passwords, pin codes, security questions, one-time passwords, hardware tokens, biometric information, voice recognition, etc. in order to make it as difficult as possible for attackers to access accounts that have high privileges. With the advent of authentication applications that create tokens with 30-second expiration timers; organizations could utilize this key-based Secure Shell (SSH) authentication instead for remote connections [125].

An extension of the multifactor authentication is quorums. Organizations must establish quorums where multiple people need ”to authorize administrative and configuration changes” [151]. The Microsoft Azure environment [184] utilizes the quorum function against rogue admin scenarios however, this is equally applicable to stopping the lateral movement and privilege escalation of ransomware within the organization. It could also block alterations to backup data.

Organizations must also adopt zero-trust access strategies. Therefore they need to continuously verify users and devices in their networks and always assume compromise [148, 151]. Moreover, an extension of the zero-trust jargon is that users must get minimum access rights and permissions, also named the principle of least privilege, where users can only obtain permissions if it is strictly necessary for operational continuity [96, 109, 177].

The access controls could also be done specifically based on roles [94]. This principle is called the ”Separation of Duty” [96] and it is important for storage security. The privileges for data management, host administration, and data protection must be distributed across different roles, and these roles must not be assigned to the same user [96]. Data management here refers to having the privileges of creating and mapping a storage volume or network share. Data protection on the other hand; refers to having the ability to configure, stop, and delete backups. Lastly, the host administration role refers to having the privilege of creating or deleting objects in the storage controller. As an example, the Microsoft Azure Backup role-based access control allows the segregation of the mentioned duties within the IT team to only allow necessary access to backups [184].

Considering the attacks ransomware do on user credentials, such as utilizing tools that can dump plaintext passwords, it could be a good idea as well to disable plaintext passwords [109]. Furthermore encrypting login credentials at rest is also a way to introduce additional security for the credentials [96].

Lastly, organizations must also restrict access to and physically secure devices and media that backups and sensitive information [146, 152].

An important aspect of security is ensuring that the backups are encrypted both at rest and in transit [146, 155]. This is especially true for data backups which hold a lot of sensitive data that needs to be compliant with regulations such as the GDPR [160]. However then, key management becomes a key consideration as the organization must ensure that the keys are only available to a selected few to ensure ransomware cannot reach it but are accessible if the necessity arises for recovery [90, 155]. This method however is only to secure the data against data exfiltration by the ransomware. This way, if the attackers get the data, they will not be able to access it. However, it is important to note that ransomware can utilize secondary encryption to encrypt backups if they get access to them.

The Australian Cyber Security Centre (ACSC) [107] recommends that organizations must segment their networks and establish policies in order to restrict traffic for remote access. This way, the lateral movement capabilities of ransomware actors will be severely limited. This is noted to be especially important in cloud and hybrid environments [159]. The NIST Framework [96] notes that the usage of different Virtual Local Area Networks (VLANs), using separate IP subnets will increase the network separation across different segments. [98] adds that correct network separation could be established by identifying critical business systems, isolating them, and applying proper network monitoring practices and network security controls. Regarding the network, [169] notes that closing all inbound ports except those required by the backup software will be critical in keeping the network safe.

CIS Controls [142] additionally note that not only network segmentation but data segmentation is good practice. The NIST Frameworks [96] explains this concept well; they note that with regards to the storage management, access, and their usage and protection must be differentiated between different data planes. NIST divides the data planes into three; the data consumption, management, and protection planes. The backups are stored only under the data protection plane which has protocols, operations, and network access policies with regard to backup processes. Furthermore, backup data needs to be separated from the primary domain where the production data is located in [96]

Monitoring has been described prior as it is an overarching concept. IT professionals must be vigilant towards indicators of compromise and must always ensure to monitor network-related configurations and access patterns [93, 161]. ICT professionals must further monitor; PowerShell execution [114], the execution of backups to see whether they are done with regards to the pre-defined policy, the completeness of the recovery copies and their good health, and the cyber hygiene of backup copies by utilizing antivirus scans, analysis, and vulnerability scanning [96]. It is important to track the errors and failures in backups and must be swiftly addressed [179]. 24/7 Security Information and Event Management (SIEM) or Security Operations Centres (SOCs) could be beneficial for the monitoring in organizations [170]. Further monitoring could be done with regards to early detection by investing in protection solutions that use behavioral analysis on ransomware [94, 104].

Patch Management is another key consideration with regard to ransomware, as they are known to identify and exploit vulnerabilities in the system to do lateral movement and escalate privileges to ensure access to backup data. Software including the backup solutions and antivirus software, operating systems, firmware, applications, and etc. must be regularly and timely patched [107, 113, 117, 153]. Organizations must also conduct regular vulnerability assessments [107].

Considering the ransomware attacks that utilize OS services, it is imperative that organizations disable or remove unnecessary services and protocols in their devices [96, 99]. For example [107] and [177] note that the Remote Desktop Protocol service must be removed if it is unnecessary. [109] and [107] explain that PowerShell must be restricted to only administrators and must be actively monitored.

Moreover, a segmented IT Landscape increases the attack surfaces of organizations. Therefore it is imperative to reduce and trim the number of applications through consolidation and simplification [178]. It is also important to decrease the mass data fragmentation if multiple solutions and storage media are considered [94]. This could be done through a unified solution such as Backup as a Service (BaaS) or a modern data security and management platform [151].

Organizations might choose to utilize backup solutions that utilize the soft delete method, which keeps the deleted backups in a cache, should there be a need to restore from them [173].

Data Reduction

The data reduction has been noted under two categories. These categories are data deduplication and data compression. Data deduplication is when the duplicates of the data are removed. More specifically [96] notes that the data deduplication replaces multiple copies of data with references to a shared copy. An example they give is that if there are 500 identical blocks, only one of them will be stored and the rest will be referenced to that copy. Data compression is noted to be commonly used by tape backups [96]. Compression must be done with the lossless method in order to not lose any crucial backup data.

[186] notes that these features can reduce the size of the data that is needed to be backed up. Combining data reduction with establishing criticality metrics to only backup the most critical data can lead to faster backup times and ensure organizations are not limited by capacity constraints.

Recovery Principles

Ransomware attacks are known to infect backups in order to establish control of the recovered system to do their attacks again. [142] and [147] note that organizations must either do their recovery with good backup data that predates the believed data of the original infection, or that systems must be fully wiped clean and rebuilt from scratch.

Here it is important to note that finding out when the original infection happened takes a long time and is not guaranteed to deliver reliable results regarding whether a backup is totally clean. Even if the backup is clean, it could be so that the restored system will have the same vulnerabilities that have been exploited to give the ransomware attackers access. This analysis downtime might not fit the recovery time objectives (RTOs).

Testing Recovery & Training

Organizations might be happy that they have their backups in place, however, the backups are still at huge risk if there are no tests to ensure they are fit-for-use during recovery. [148] notes that backup recovery processes must be regularly tested. If they are not, then organizations will not be aware if their backups are corrupted or not [154]. [107] notes that ”a robust backup and disaster recovery plan that is well exercised will give governments, businesses and individuals greater power when making a decision about recovering of their files should they fall victim.”

There are certain ways that the tests must be made. First, when a backup is made either by automatic or manual means, the backups must be verified [149]. Second, the backups must be restored and this restoration process must be documented at least per year [144, 149]. The CIS Controls on the other hand These tests must also be benchmarked with respect to the organizational recovery time objectives (RTOs) [96].

The testing however is also related to end-to-end testing [153] to exercise the backup & recovery protocols in order to train employees so that they are aware and are ready for a cyber attack situation. [98, 163] notes that regular training sessions must be made for improving the staff competence regarding recovery.

Utilizing Software as a Service (SaaS) and Backup as a Service (BaaS)

Many organizations utilize Software as a Service(SaaS) products such as Microsoft Office 365. One important thing is however to ensure that the files generated through these services need to be backed up properly. [181] notes that many SaaS products are utilizing cloud-to-cloud (c2c) data backups which replicate a backup made in the cloud on another for increased security. [169] notes that SaaS backups could be utilized as an alternative to on-premise backup for organizations.

In addition to this, many organizations who offer SaaS solutions, also promote Backup as a Service (BaaS) solutions so that the customers always have their critical data available to be restored [188]. The main players in the $10 billion BaaS market are Veeam, Datto, Commvault, Cohesity, Acronis, Backupify, SolarWinds and Spanning, APEX Dell Technologies [188].

A Cloud-Based Backup as a Service offers capabilities to collect, reduce and encrypt data that comes from the SaaS applications and to transfer these data to cloud-based storage with respect to the established backup plans [188]. Some BaaS solutions also allow for hybrid backup environments, which can synchronize on-premise and cloud backup storage. The BaaS solutions can also verify backups once they are automatically created [180]. [188] notes that smaller to medium size companies can leverage the power of BaaS software to compensate for their IT staff limitations by allowing a third party to do the backups for them. [168] notes that the BlueXP backup and recovery for example allows specific protection against ransomware which also scans the backups in order to monitor any alterations made to the files. If there are any changes, the IT professionals get alerted.

The SaaS and BaaS solutions thus might reduce the challenges faced regarding the IT overhead and costs of backup & recovery.

Collaboration

The NIST recommends collaboration between organizations as they promote information exchange to mitigate cybersecurity risks such as ransomware. The information-sharing arrangements have also become a voluntary part of the new DORA regulations.

Information-sharing arrangements are nothing new. There are many entities known as Information Sharing and Analysis Centers (ISACs) where organizations come together sector-wise in order to share information on: indicators of compromise, new vulnerabilities, and attacks that they see. An example could be the Financial Institutions Information Sharing and Analysis Centre (FI-ISAC) which has been established by the Betaalvereniging Nederland (Payment Association Netherlands) in 2003. This institution has been linked to the National Cyber Security Centre, which is under Ministry of Justice & Security, in 2014.

3.4. Backup & Recovery Challenges

In this section the challenges that organizations face while generating and maintaining backups, and recovery are revealed. Out of 102 files 42 of them give insight into the challenges. Many challenges are faced due to the limitations of the backup storage media, the lack of integrity in backups, inadequate organizational security and backup policies, the lack of testing of the backups and recovery, and insufficient resources, high costs, and lack of knowledge. The codebook that has been generated through qualitative content analysis, followed by the abstraction figure generated which will create the narrative structure of the section is given in Table 3.8. The ”#” column on the codebook represents the number of times a particular concept was mentioned and is cumulative for top-level categories. The codebook under this section answers the third subquestion. In-depth explanations of the challenges are made starting from Section 3.4.1.

Table 3.8: Codebook of Challenges in Backups and Recovery

Category/Subcategory/Properties			#	Citations
Backup Storage Media Challenges			11	[93, 106, 139, 156, 174, 179, 183, 188, 193, 194, 195]
	Cloud Limitations		6	[93, 139, 156, 183, 188, 194]
	Disk Limitations		1	[179]
	NAS Limitations		1	[193]
	Tape Backup Inadaquacies		4	[106, 174, 179, 195]
Complexity Management			24	[90, 93, 94, 97, 149, 151, 157, 170, 174, 179, 180, 182, 183, 186, 187, 188, 195, 196, 197, 198, 199, 200, 201, 202]
	Coordination and Communication		6	[151, 179, 195, 196, 197, 202]
		Geographically Dispersed Data	1	[179]
		Incompatible Solutions	2	[151, 195]
	IT Overhead		10	[180, 201] , [94, 151, 157, 179, 196, 198, 200, 202]
		Complexity in Data Virtualization	2	[180, 201]
		Legacy Backup and Recovery Solutions	4	[151, 198, 200, 202]
		Tiering Overhead	1	[179]
	Regulatory and Provider Challenges		11	[90, 93, 97, 149, 174, 182, 183, 188, 199, 201]
		Outflow of Sensitive Data	3	[90, 149, 174]
		Outside Region Regulatory Requirements	1	[97]
		Regulatory Complience	2	[93, 199]
		Slow Reaction from Provider	1	[90]

Table 3.8: Codebook of Challenges in Backups and Recovery

Category/Subcategory/Properties			#	Citations
Inadaquate Security and Backup Policy			20	[90, 93, 98, 99, 107, 148, 151, 154, 156, 162, 170, 173, 175, 179, 182, 186, 196, 197, 198, 203]
		Access Management	8	[90, 98, 99, 107, 148, 162, 175, 182]
		Insufficient Monitoring	2	[90, 179]
		Lack of Documentation of Backup and Recovery Policy	6	[90, 170, 179, 182, 186, 196]
		Unpatched Software and Vulnerabilities	4	[93, 151, 173, 197]
Lack of, Corrupt, Infected, Incomplete and Unaccessible Backups			14	[90, 96, 148, 154, 157, 158, 175, 179, 182, 186, 188, 198, 199, 202]
	Backup Solution Failure		2	[154, 158]
	Lack of and Incorrect Configuration of Backups		10	[90, 96, 148, 154, 157, 175, 179, 188, 198, 199]
		Incompatible Backup Type	3	[154, 175, 179]
		Low Frequency	3	[96, 188, 199]
		Low Retention Period	1	[96]
		Loss of Cryptographic Keys	1	[90]
Resources, Costs and Knowledge			19	[90, 93, 96, 150, 151, 154, 157, 158, 165, 169, 172, 173, 175, 179, 182, 193, 195, 200, 201, 202]
		Data Backup Speed	5	[90, 150, 154, 158, 195]
		Data Size and Insufficient Capacity	11	[90, 151, 157, 165, 175, 179, 193, 195, 200, 201]
		Financial Costs, Budget and Willingness to Pay	4	[93, 172, 179, 202]

Table 3.8: Codebook of Challenges in Backups and Recovery

Category/Subcategory/Properties			#	Citations
		Lack of Cybersecurity Training, Awareness and Experience	3	[93, 151, 169]
		Lack of Knowledge of Write-order fidelity (Sequentiality)	1	[96]
		Insufficient Maintenance	2	[90, 182]
Testing			9	[90, 150, 157, 175, 179, 182, 186, 199, 202]

The table will again be analyzed by focusing on the largely focused concepts to understand what are the main trends and mentions in the material. The challenges codebook shows that the main focus is on; complexity management (24 mentions), inadaquate security and backup policies (21 mentions), lack of resources high costs and knowledge (18 mentions), the lack of or the corrupt, infected, incomplete and unaccessible backups (12 mentions), the challenges faced with backup storage media (11 mentions) and testing (9 mentions).

Under complexity management, the most notable categories of challenges are IT Overhead (11 mentions) and regulatory or provider challenges (11 mentions). Inadequate security and backup policy on the other hand mainly focuses on access management (8 mentions) and the lack of documentation of backup and recovery policies (6 mentions).

Under resources, costs, and knowledge mainly; data size & insufficient capacity (9 mentions), data backup speed (5 mentions) and high financial costs, low cybersecurity budgets, and the willingness to pay (4 mentions) are mentioned. The challenges faced with backup storage media are focused on the cloud (6 mentions) and tape backups (4 mentions). An abstraction of the categories is given in Figure 3.3.

The main categories that the material has been classified under are; Backup Storage Media challenges, Backup Integrity Challenges, Complexity Management, Inadequate Security and Backup Policies, Testing, and Resources, Cost, and Knowledge. The following sections delve into each category respectively.

3.4.1. Backup Storage Media Challenges

Organizations have many choices when it comes to choosing a backup storage media. Each choice, however, comes with their own benefits and limitations. The material that was directly related to ransomware attacks in the analyzed material were; Cloud, Disk Storage, NAS, and Tape backup limitations.

Cloud Limitations

The cloud offers huge cost benefits for organizational storage. It also comes with an additional layer of protection as it is off the main network and is sustained by a third party. However, these benefits also may become their main drawbacks.

The first notable limitation is the data restoration speed, as it is primarily linked to bandwidth limitations, which can lead to slow data recovery from the cloud during a crisis [188]. The cloud providers usually work on the pay-as-you-go paradigm. The

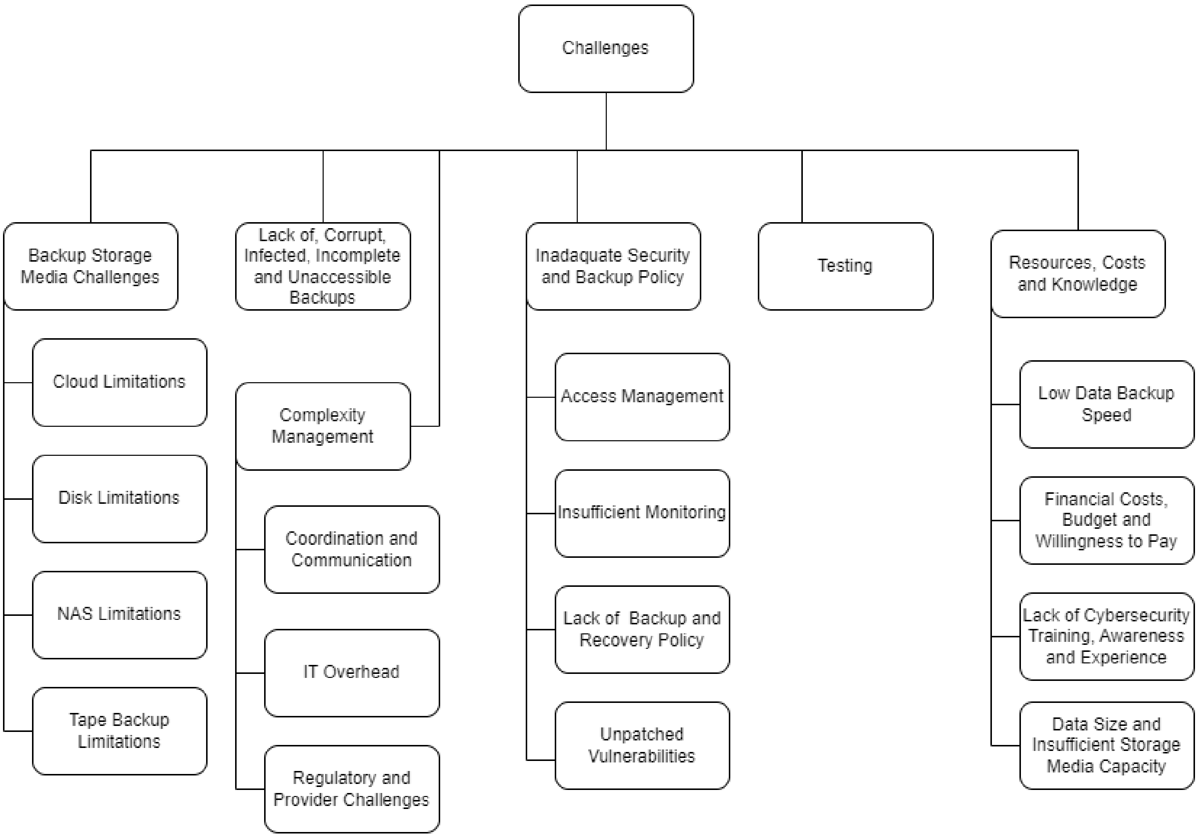

Challenges
Backup Storage Media Challenges
Cloud Limitations
Disk Limitations
NAS Limitations
Tape Backup Limitations
Lack of, Corrupt, Infected, Incomplete and Unaccessible Backups
Complexity Management
Coordination and Communication
IT Overhead
Regulatory and Provider Challenges
Inadaquate Security and Backup Policy
Access Management
Insufficient Monitoring
Lack of Backup and Recovery Policy
Unpatched Vulnerabilities
Testing
Resources, Costs and Knowledge
Low Data Backup Speed
Financial Costs, Budget and Willingness to Pay
Lack of Cybersecurity Training, Awareness and Experience
Data Size and Insufficient Storage Media Capacity

Figure 3.3: Abstraction for the Categories of Challenges

high bandwidth consumption and the cost of storing and recalling cloud data may necessitate local storage or other considerations regarding backups. The scalability is then a challenge for organizations who might have increased storage needs [194]. Moreover, many third-party providers who offer public cloud services often utilize the shared infrastructure. When data from multiple customers are stored on the same servers, there will be privacy risks involved. The infrastructure will also then might suffer with respect to reaching the organizational RTO goals and will not provide a smooth data recall and recovery.

A cloud service is basically connecting to a third-party computer and servers for storage [183]. Thus, if the cloud provider is affected by a disaster such as a fire or a data breach, this constitutes a big challenge for organizations. The OVHcloud case [183] shows this issue quite well. Due to a regional fire, customers who had not purchased a higher plan which allowed their data to be backed up also to another region lost all their data in the cloud.

Data protection and vendor transparency is another important consideration. Some cloud-storage vendors clearly state their backup and recovery processes, including the testing and the protection of the data. They also might be transparent on the division of responsibilities between the customer and themselves during the purchasing contract-making stage. However, not all cloud-storage providers are this transparent [183]. In a case of a ransomware attack, it is then necessary to have assurances that are documented about backup and recovery from the third party.

Turning to the protection and the security of the data, there are concerns for data that is in-flight and at rest and complying with regulations across all data [194]. Quite related, [93] notes that some cloud providers also do not have encryption protocols for the data, which poses a potential security risk. Ransomware who have access to such data and the backups of these data can freely exfiltrate it in such a case.

The cloud-syncing services such as Dropbox, OneDrive, and SharePoint, or Google Drive must not be used as a sole backup method. They introduce the risk of automatic synchronization of the ransomware-infected or encrypted files. This then would play into the hands of the attackers in order to destroy backups on the offsite storage that is the cloud.

Disk Storage Limitations

The downsides of disk storage, according to [179], are that it grows to be more expensive than tapes and that this type of data might not be suitable for archiving, where the mobility of the storage media is especially important. The disks carry risks due to their physical nature, as they might be susceptible to environmental factors. It also becomes a single point of failure should the backups not be stored on an alternative storage. These directly relate to the difficulties of keeping the recovery data safe and that the organizational costs do not outweigh the benefits of backups.

NAS Limitations

[193] notes certain NAS limitations these are; huge data sizes that NAS cannot keep up with, NAS backups being incompatible between vendors, having limited compatibility with database applications, and that the Network Data Management Protocol allowing only nine incremental backups before a full one is made. The exponential growth in data volumes since the NAS backups were first established has currently reached petabyte levels in organizations and thus outscale the capabilities of NAS.

Moreover, NAS devices have an issue of interoperability between those belonging to different vendors, as the backups made from one device cannot be seamlessly restored to another due to operating systems. To circumvent this issue, the Network Data Management Protocol (NDMP) was introduced in 1966, which allowed the backup servers to communicate with NAS through a separate management layer. However, the NDMP is still lacking in providing true interoperability due to the lack of data format standardization which results in vendor lock-in. The NAS backups also have limited compatibility with database applications. Lastly, the NDMP's limitations on incremental backups which has originally been designed to accommodate tape-based backups, have become unnecessary compared to utilizing disk-based storage.

Based on the challenges if NAS cannot handle the sheer size of the data, this would lead to capacity issues which may conflict with the retention period of the backups, and overwrites may be necessary. Incompatibility between vendors and databases makes it hard to store the backups & to recover from them should the need arise. The NDMP limitation on incremental backups, combined with the amount of data that needs to be stored then, will put pressure on the data capacity and thus add to the financial costs or increase complexity.

Tape Backup Limitations

The tape backups present several issues that concern data integrity, performance, and recovery. The performance here is also directly linked to the speed of recovery.

Regarding data integrity, data stored on tapes degrade over time, risking damage or loss [195]. The tapes are also more susceptible to environmental damage and human errors such as loss and unintentional damage. The losing of the tapes is a risk, especially when they are moved offsite [179], and is a critical risk if they contain sensitive data.

Moreover, tapes are not compatible with incremental backups due to their need to maintain a constant write speed. If the incoming transfer rate of backups is slower than the speed of the tape write speed, it will have severe performance issues as it has to stop, reposition, and start again [174]. This could be circumvented, however, by putting the backups on a disk, creating a full backup, and then transferring the entirety of it to a tape at high speeds [174]. Recovering large amounts of data and system backups, however, due to the limited read and write speeds on the tapes, could be slow and complicated [106]. The relative slowness of recovery from tape versus a disk-based restore can be attributed to the sequential access method that the tapes utilize [179]. The sequential access means that the tapes access data one record at a time from the beginning to the end.

3.4.2. Backup Integrity Challenges

Another big challenge that organizations face during a crisis aftermath is the lack of, corrupted, damaged, infected, incomplete, incompatible, or unaccessible backups. There are a plethora of reasons this might be the case.

A reason for corrupted or damaged backups could be that the backup solution had failed [154, 158] during the backup process and that the files were not properly validated afterward. This usually happens due to over-reliance on backup and recovery solutions [202]. There is also a chance that some of the data on the storage of the organization are inconsistent due to using outdated solutions [154]. Furthermore, the manual backups are quite unreliable [157] as human error might come into play. Backup schedules may stumble over timeliness as organizations may not keep to the schedule.

Another could be that the backup type has been selected wrong [154, 157, 175] and perhaps is incompatible. There are a lot of backup types that complement different types of data. For example, some backup types are more suitable for file-level backups, whereas others upload a full snapshot of the machine.

Another issue is that backups for the affected systems or data might not be there at all [154]. [198] notes that 14% of all data is not backed up. If what is to be backed up in order to bring the infrastructure back is not decided upon correctly, it leaves certain data portions at risk. [188] states that 60% of companies do not conduct daily backups. [96] and [199] notes that if there are not enough backups generated or retained, at least some portion of the data will be lost. [186] reminds that the excessive time to recover, partial recovery, or failure of recovery are usually due to missing and incomplete [199] prior backups.

Regarding the backup inaccessibility, it could be that the backups were encrypted prior to establish security. However, in this case if proper key management is not done, the cryptographic keys might be lost [90].

Lastly, a challenge is that the backups might have been infected with ransomware. This increases the time for recovery as the organizations must then analyze and clean their backup files. This is exasperated by the fact that a forensics recovery team and other cybersecurity professionals may need to be involved to fix the backups and to make sure the newly restored environment will be clean.

All these challenges relate to the integrity of the backup files, which may downgrade recovery performance.

3.4.3. Complexity Management

Organizations must be able to manage complexity with regard to backup and recovery. The main issues to note are the IT overhead, regulatory and provider challenges, and the coordination and communication among different teams.

IT Overhead

Organizations require quite a lot of tools to establish their core value offering. However, the integration of new software and applications presents significant challenges with regard to introducing new complexity and uncertainties in backups [157]. This situation is also caused by the dependency on legacy IT systems, which might not be enough to handle threats such as ransomware [198]. [196] identifies that the IT ecosystem in organizations and supply chains demands an integrated approach to recovery.

The varying IT Landscape and its sheer size expand the attack surface of organizations for ransomware attacks. These attack surfaces increase as the multiple end point products that target backup create silos that span both on-premise and multi-cloud environments [94]. The data virtualization that many organizations utilize, and the involvement of new devices such as mobile phones complicate the data recovery processes where the virtual environments and mobile devices, etc. present new and unique challenges that do not exist in traditional computing environments [201].

The proliferation of backup and recovery tools, mixed with the increased data fragmentation that originates from the use of cloud providers and Software as a Service (SaaS) applications; make it difficult also to implement and manage data protection on backups [180]. If the existing solutions are not well integrated, this will create a never-ending complexity [151].

The legacy backup and recovery solutions have also been noted to create barriers to data security [202]. As the legacy backup solutions were not designed with the newer solutions in mind, there could be problems in the migration of the backup data [195]. [151] notes that the legacy systems that create the backups require additional IT specialists to be effectively operated. The legacy sotrage media can also make it challenging to scale up backup storage systems and to migrate towards modern infrastructures [200].

Lastly, backup data requires different levels of protection with regard to the criticality of the data in question. These backups also require their corresponding established backup schedules and retention policies, which is made difficult as the data volumes grow exponentially and the number of servers to manage increase [179].

These challenges regarding the IT overhead relate to organizations starting to overlook things, increasing the attack surface. The establishment of backups and recovery becomes more arduous and the attack surface is then targeted by ransomware groups.

Regulatory and Provider Challenges

The regulatory and provider challenges are related to mostly the outflow of sensitive data, regulatory compliance, and the possible slow reaction from providers during a crisis.

With a third party arrangement, the data's control is on their hand. If there is a ransomware attack, the data recovery is also then related to the service vendor. However, [201] notes that many vendors do not let the detailed configurations of their products. This may lead to data recovery being harder.

It is also the responsibility of the provider to protect the backup data that is saved on their platforms. The Cloud Hosting, Infrastructure as a Service (IaaS), Platform as a Service (PaaS), or Software as a Service (SaaS) solutions are noted to be not protected sufficiently [182]. [93] notes that third parties can become an attack vector for organizations through a supply chain attack. Thus, third parties may create attack chances for ransomware.

[183] notes that there must be transparency of how the service provider stores the backup data, and how they are protected. It is imperative that the backups are protected with respect to the 3-2-1 strategy to ensure their safety against ransomware attacks. [183] notes that many cloud contracts do not mention backup or talk about disaster recovery, and it is stated that if they do, the responsibility of the backups and disaster recovery are made the responsibility of the organization. This is quite crucial as an organization does not demand proper answers regarding backup and recovery and these are overlooked then they face a situation where they have responsibility and accountability but no control over their data which severely limits their recovery capabilities.

The provider limits then are also quite important to note. For example, Google notes that ”but that is not a guarantee that your data is protected”, or that Microsoft states that ”all online services suffer occasional disruptions and outages, and Microsoft is not liable for any disruption or loss you may suffer as a result.” [188]. These must remind organizations that the availability of their backups to them are not guaranteed even during a crisis. [90] supports this claim as that unfavorable contractual conditions may result in backups being available at short notice. This then would have a negative effect on the downtime.

It is also important that the cloud storage providers are complying with the necessary regulations. [90] notes that the outsourcing of the backups may create sensitive information outflow, which is protected by mandates. [199] states that GDPR, CCPA and HIPAA regulations have strict guidelines on data handling, which include backup and recovery. [97] notes that the storage of backup copies with external providers who offer cross-regional services might be also limited due to regulatory requirements to house data within certain geographical regions. This decreases the extra security on the backup data and increases the amount of overhead in organizations so they need to plan all these efficiently for a proper backup and data recovery mechanism.

Coordination and Communication

Coordination and Communication are critical, especially during a recovery scenario. [202] states that there is often misalignment between IT and security teams. [151] states that IT and security roles are siloed, and in their survey, they have found that 31% of the Security Operations (SecOps) respondents believed the collaboration with

IT is not strong. Moreover in a cyber attack scenario more often than not internal and external communications are restricted due to reliance on network-enabled channels [196]. Not having alignment, mixed with the fact that communication is limited during a crisis will dampen the recovery efforts, especially if there are no documented and tested recovery plans from backups.

[197] states that the restoration from backups must be well coordinated across ”backup and storage administrators, application DBAs, developers, and networking personnel.” This is quite crucial as even if files or virtual machines are recovered by backups the people responsible of the IT infrastructure need to be able to decide and act together to start the application and database servers.

Another factor to consider is that the backup data might also be too geographically dispersed, in places that do not have a dedicated IT staff on-site [179], which would need extra monitoring, and alignment by IT professionals in order to ensure their availability during recovery.

3.4.4. Inadaquate Security and Backup Policies

Organizations might have longer downtimes during recovery due to inadequate security and backup policies [90]. They might have poor access management, lack of documentation, insufficient monitoring, and unpatched software and vulnerabilities.

Other than the main 4 issues that will be described in the following sub-sections, [90] notes that organizations may find challenges when the backups are not being stored in a secure location. [156] notes that even one of the most popular solutions, Amazon Web Services, does not have separation or airgap between the primary data and the snapshots generated when the account that accesses the primary data are the same.

Access Management

Access management is one of the most important pillars of defense against ransomware attacks. If this is not properly established and monitored, then the ransomware actors can move laterally and escalate their privileges quite freely and destroy the backups in the system. [90] notes that if access rights are granted generously or that passwords are easy to guess the attackers might cause more security incidents.

[182] gives insight into what could go wrong regarding access management prior to recovery. To restore to a fresh infrastructure the domain network server (DNS) needs to be set up and the recovery software VMware, Hyper-V, etc. rely on the DNS working. [182] notes that during setting up the DNS, the teams might have trouble as; they don’t know the local administrator passwords, which is a requirement when the active directory is offline, or that the directory services recovery mode (DSRM) administrator password is also not known which can recover the active directory. Moreover, the password vaults may be part of the production environment and might be unavailable. The teams might also not know the credentials for core services. [154] also notes that a reason that backups are inaccessible could be due to the loss of access to the backup storage as the credentials are lost.

The prior paragraphs show how if access management is not monitored well, it makes it easy for attackers to get access and destroy backups. If the backups are not properly stored, it makes downtime during a crisis longer.

Lack of Documentation

The success of the normal backup and restore procedures to continue business operations depend heavily on the documented plans prior to a disaster [179]. However, the lack of documentation on the backup, recovery, and in general disaster recovery and incident response are one of the most critical pain points during recovery.

[182] notes that these documentations are often not available during the downtime of primary systems and that the process to recovery is not fully documented. [186] state that in their research two out of three respondents had inadequate processes and had no test documentation. FOX-IT [170] implies that University Maastricht wanted to map out their IT inventories after the attack and that they had insufficient understanding of the number of (in)active computer and server systems in their domain.

These show that in reality, even if the companies may think that they have their processes documented quite well, there is still a lack of documentation and specifying the characteristics of plans that include backup and recovery.

Insufficient Monitoring

Monitoring by IT Specialists is quite important. [90] states that software that writes the backups may overwrite the backups which still need to be retained. If the IT specialists are not informed about this overwriting or whether the verification tests [179] done are successful, then this could lead to unusable backups during recovery. Thus the maintenance of the backup and disaster recovery systems is quite important. They need to be maintained with errors being fixed at the right time. Errors in the backups will lead to failed recoveries.

The monitoring also quite relates to whether operational plans regarding backup and recovery are done correctly and that everything is executed, verified and tested well.

Unpatched Software and Vulnerabilities

It is already established that ransomware uses the unpatched vulnerabilities that they can find in the devices connected to the organization's network. [173] note that 42% of vulnerabilities are exploited after a patch had already been released by the software vendor to fix it. Unpatched vulnerabilities allow ransomware to get more access and move laterally within the system. This relates directly to poor security of the backups and the probability of them being compromised.

The problem here is that patching software takes time and is costly [151], the IT Landscape is too complex [93, 197] and the backup solutions have legacy security models.

3.4.5. Testing

Not enough testing is a prevalent issue in organizations. Many organizations do not test their backups to see whether they are fit for purpose. Moreover they also do not test their cyber recovery, disaster recovery [182], and incident response plans. The lack of end-to-end testing may provide challenges in a real-life scenario where the responsibilities etc. are not properly exercised.

[202] found in their research that only 54% out of 1625 IT and security leaders had tested their backup and recovery. Without proper testing, there is no knowing if the data backups will allow for full recovery [90, 157]. Veeam's 2016 Availability Report

[186] shows that less than 50% of respondents test backups only once a month, and among those who test, 74% test less than 5% of their total backup quantity. There is further support for these numbers in [150].

3.4.6. Resources, Cost and Knowledge

Resources, Costs, and Knowledge play a huge role in decision-making regarding tradeoffs to navigate the organizational complexities of managing backup and recovery. Main issues are noted to be: data size & insufficient storage capacity, low data backup speed, financial costs, low budget & willingness to pay, the lack of cybersecurity training, awareness, and experience, and lastly, lack of knowledge in sequentiality in recovery.

Data Size & Insufficient Storage Capacity

[90, 151, 157, 180, 201] note that organizations face data storage limitations due to the exponential growth of data, also named the data sprawl. [165] notes that the data size contributes to recovery spanning days which tanks the operational productivity of an organization.

[179] adds the fact that with the increase in server amount and data sizes the likelihood for a failure also grows in backup operations, and becomes harder to monitor.

Following this [90] notes that organizations also often neglect purchasing storage media with enough capacity for regular backup operations.

Data Backup Speed

Organizations face issues when it comes to data backup speed and data recovery speed. Rationally, the data that can be transferred is based on the bandwidth of the network or the bandwidth of the backup solution.

[154] and [158] note that backups might take very long to finish which slows down the RTO and RPOs. Slow internet connection and network performance issues can contribute to this problem. [90] notes that the time needed to perform backups are also increasing due to the data size. They note that in the worst case if a new backup begins prior to the previous one ending due to the speed, then no new backups might be made which could also lead to the previous one being terminated. [195] brings highlight the fact that the legacy backup solutions exasperate this problem as they have longer prolonged backup windows and that backup processes become unmanageable.

[150] found in their survey that 38% of organizations switch away from their backup solutions due to them being slow. They further found that about 76% IT professionals report experiencing bandwidth constraints during backup processes. 62% of the commonly accesses reports are related to the backup performance/speed/throughput [150].

Thus, low backup speeds contribute to the increased complexity of managing them and relate to not meeting RTO and RPO goals.

Financial Costs, Budget and Willingness to Pay

It could be seen that each backup strategy comes with a price tag attached to it. An example is given by [179], where they note that the costs of offsite storage vary and add up with regards to how long the tapes are stored, the length of time, etc. Similar considerations are made in cloud environments where the primary business is a pay-as-you-go approach.

However, compared to the costs, organizations do lack the resources, or the allocation of it for the cybersecurity budget and fall behind [93, 173]. [202] states that in their survey ”47 percent of IT and security leaders believe their 2023 budgets is not enough of an investment, while 27 percent expect their IT and cybersecurity budgets to decrease this year.”

There is also a willingness to pay in organizations as [172] show in their report that there are many organizations, about 46% in their questionnaire, who choose to pay the ransom instead of utilizing backups.

Lack of Cybersecurity Training, Awareness and Experience

[93, 151, 169] all note that organizations do not properly educate and train their employees due to a lack of time and resources. Cyber security awareness then is a must, and the hackers very well understand that the backup servers are quite under-protected and are usually administered by junior personnel who have less experience in information security [169].

The lack of awareness and education can lead to successful phishing attacks by ransomware to get initial access to the network moreover, if the backup and recovery personnel are not well prepared against cyber attacks, ransomware actors can just exploit the gap in knowledge and find ways to destroy backups to make recovery impossible.

Lack of Knowledge of Write-order Fidelity (Sequentiality)

When bringing operations back from backups one important thing is the write-order fidelity, which is closely tied to sequentiality. Write-order fidelity refers to preserving the original sequence in which the updates or changes are applied to a dataset. Sequentiality means that the data should be written and restored in the same order it was written.

If the backups do not preserve this sequentiality the system then will be restored to an inconsistent state [96]. This will then hamper the recovery and cause excess downtime to operations. The knowledge of this must definitely be known by the organization’s IT teams.

3.5. Conclusion

This chapter has focused on the utilization of the qualitative content analysis method over 102 sources that also include 11 cybersecurity frameworks, noted as best and established practices. The remaining 91 sources are collected from backup software and antivirus solution providers, and cybersecurity blogs. The inductive nature of the research has allowed for a generation of codebooks that identify concepts that answer to all three subquestions that are related to ransomware attacks, backup & recovery strategies, and challenges faced by organizations. The fourth subquestion which denotes the implications of the discovered findings with respect to providing a recommendation list is partially answered through generating the basis of argument for Chapter 5. The main findings of this chapter are again categorized in the order of the subquestions for the purposes of generating a conclusion.

With respect to the first subquestion ”How do the most salient ransomware families target and compromise backups?” the main findings were that ransomware: attack the backup integrity in the storage management systems. To this end they do reconnaissance of the system through utilizing legitimate network scouting tools in order to identify file shares, account privilages, and etc. Ransomware actors then corrupt, delete, encrypt and infect backup copies due to their high privileges in the system that they have acquired through continuous privilege escalation and lateral movement following the reconnaissance. The ransomware tend to identify and terminate services and processes that can create, modify, and recover from backups. Ransomware tend to also utilize vulnerabilities on backup software in order to be able to acquire access to backup & recovery specific services and storage. The ransomware also commonly use legitimate operating system utilities that are originally reserved for system administrators; to delete shadow copies, and to deactivate recovery environments. Lastly, the ransomware terminate antivirus software and other utilities in the operating system that can be used to mitigate the impact of the virus.

The findings for the second subquestion, namely, ”What are the current state-of-the-art backup & recovery practices that are (or can be) employed by companies to safeguard their data and ensure operational continuity?” are that: Organizations must ensure strong backup and recovery policies. These policies must entail best practices with respect to backup generation, backup protection, recovery testing, recovery principles, and collaboration. Furthermore organizations can utilize SaaS and BaaS software to decrease their IT overhead and spread out their backups with accountable third-parties. Data reduction methods can achieve less data sizes reducing the storage capacity necessity, and the recovery principles state that it is always good to wipe and rebuild from good backups. Collaboration is key and can be essential to keep up to date with recent developments in ransomware attacks with the help of knowledge sharing between organizations.

Here for backup generation what is most important is to regularly create backups including redundant backups, that have been made cyber attack resilient through isolating the recovery data with airgapping practices, moving the backups offsite or offline arrangements. These backups could also be made immutable for further protection by utilizing methods such as the write once read many, etc.

For protection on the other hand, it is essential that good access and privilege management is established in the organization. In the case of data exfiltration it is important that all sensitive data including the backup data must have been encrypted. The organizations must also ensure that they always monitor networks, backup data, and other application health that can be exploited to run a supply chain attack. Further protection can be realized through network and data plane segmentation and seperation. To ensure that vulnerabilities are not available for attackers patch management is also a key factor for organizations. Furthermore, organizations must disable all unnecessary services and protocols, streamline their IT landscape, utilize soft delete for strong protection.

Backup and recovery policies of organizations must include information on the content, criticality, prioritization and the characteristics of the backups. It must also include the inventory of backup storage and how the lifecycle of the data are managed. They must also establish operational plans regarding the saving of the data. There must also be scheduling of periodic audits. Moreover, responsibility and accountability for during a crisis must be well established, including the communication and cooperation plans for during recovery. Organizations can also select from a plethora of backup storage options which correspond with their business plans and goals where each have their own tradeoffs.

Lastly, the findings regarding the third subquestion which notes ”What are the main challenges faced by organizations regarding backups & recovery?” are mainly categorized under: Backup storage media challenges, complexity management, inadaquate security and backup policy, inadaquate testing, and the lack of resources, costs and knowledge. Furthermore, organizations are noted to face the lack of, corrupted, infected, incomplete or unaccesible backups in the event of a ransomware attack which become critical challenges for recovery and increase the downtime.

4
Semi-Structured Expert Interviews

10 Semi-Structured Expert Interviews have been conducted in order to be able to triangulate the findings of the qualitative content study to increase the trustworthiness and the validity of the prior findings through understanding if what experts say actually match with the results that have been found before. Moreover, the interviews are expected to give new insights into how experts perceive the ransomware threat regarding backups and recovery in their day-to-day work, and perhaps can give concrete examples on how decision-making in organizations is conducted on this topic to give a better understanding of why they might face certain challenges over others, etc. They also highlight certain aspects that have not been in scope prior e.g. forensics and legal matters create more downtime for organizations and thus is a prevalant challenge to ensure that regulatory and legal necessities are met in shortest amount of time possible. Moreover it is also highlighted how red teaming and penetration testing for organizational compliance plays a big role in securing organizations, but they are bound to limitations and specific actions within their playbooks that are prior established and might miss how a real advanced persistent threat might evolve in the network.

The chapter follows the already established pattern of presenting the description of the method and choices made, continuing with insights on: ransomware attacks against backups, backup & recovery practices, and challenges faced by organizations. Section 4.2 answers the first subquestion *"How do most salient ransomware families target and compromise backups?"*. In order to have an explicit understanding into the subquestion specific questions are asked. The interview protocol can be found in Appendix B. Due to the semi-structured nature, in some interviews some questions have taken priority over others and this is true for all optional questions under the interview protocol. The interviewees were asked their knowledge about ransomware attacks directed at backups. Albeit being only a single question, cybersecurity professionals who are more immersed in cyberattacks and who have knowledge on red teaming, penetration testing in organizational networks have given extensive answers. They have mainly generated input on how they percieve the ransomware attacks to be in the current era, and how they move laterally within systems, and identify and attack backups to destroy the chance of a proper recovery. They also give insight into what the known playbooks for typical ransomware are, give an understanding on the mentality of an attacker, and how the backup management software or different types of backup solutions could be attacked.

Section 4.3 aims to answer the second subquestion: *"What are the current state-of-the-art backup & recovery practices that are (or can be) employed by companies to safeguard their data and ensure operational continuity?"*. The interviewees have been asked questions about what type of backup strategies that they have worked with in organizations. They have been requested to elaborate on the complexities of data backups and their recovery, and how organizations do backups in-house or subcontracting the work to third parties. More information was requested with regards to if they have utilized specific backup solutions i.e. Microsoft Azure etc., and they have been further inquired about how organizations keep their backups safe, whether they follow specific cybersecurity frameworks as guidelines, whether they have criticality levels established for their data for more efficient backup and recovery, how often, what type of backups are created and why. There were further questions about the testing of backups, their retention periods and their creation schedules. Manyof these questions were optional and have guided the interviews in order to extract knowledge on what are the perceived best practices that people who are in the industry use and thus, answers the second subquestion through a different perspective.

Lastly, Section 4.4 answers the third subquestion: *”What are the main challenges faced by organizations regarding backups & recovery”*. The interview questions on this end were more general and exploratory. The interviewees have elaborated on what the bottlenecks they perceive are in a recovery setting in the aftermath of a cyber attack. They were also asked on what the main challenges are in backups & recovery in organizations, and how do the size of the organization, the complexity of the IT infrastructure, or other industry specific requirements impact recovery. Moreover, the interview questions have also delved into whether the unavailability or unusable backups due to many factors is a big enough perceived threat by experts. The interviewees have been also asked to elaborate on technical and operational challenges in this context and what their critical lessons learned in their prior experiences were for organizations who had failed due to facing challenges. Lastly, the experts have been asked what complexities and difficulties may arise for organizations if they follow best practices that have been laid out in the previous chapters.

4.1. Description of the Method

The semi-structured interviews are chosen in order to both validate the content analysis findings, and to explore expert knowledge in the domain of backups & recovery against ransomware attacks. The reason it is semi-structured rather than a full interview or a questionnaire is that the semi-structured interviews allow for further exploration and that the interviewees are not boxed to specific questions. This is important as the experts had very different backgrounds and that some had more knowledge on ransomware attacks, some had more knowledge on backups and recovery. Here it was important to bridge the gap by following up on questions respective to the expert’s knowledge. The questions have been separated on their corresponding sections which answer different subquestions as was described in the previous paragraphs.

The procedure of the semi-structured interviews were to first find interviewees. The interviewees were contacted with respect to a purposive sampling strategy. The main contact links were emails, and LinkedIn. Although most people did not respond, 10 were available. Then an email was sent to plan the meetings. In a follow-up email, the informed consent form and the semi-structured interview questions have been sent. All interviews happened online by utilizing the apps Microsoft Teams or Zoom. The recordings have been captured through teams and the Microsoft Word transcription utility was utilized. Due to the low quality of the transcription results, an edited transcription was done by hand. Following this, in order to fulfil the ethics requirements (described in Section 1.7.2) a technical summary that has removed all personal data and personally identifiable information. Due to the niche properties of the topic at hand, many instances or case studies have been generalized. These technical summaries were then saved in a docx file which were analyzed through NViVo qualitative data analysis software.

The data analysis of the semi-structured interviews have followed the qualitative content analysis method as described before. The three stages; preperation, organization and reporting have been carried out in an identical fashion. The analysis was again conducted inductively which generated a codebook. The aim here is to synthesize both codebooks from the qualitative content analysis chapter from prior and the current chapter in order to deliver a final recommendations list, and thus to come up with a merged codebook that can be utilized in also future research.

4.2. Ransomware Attacks

Ransomware has evolved into more modern tactics which are conducted by organized crime groups or state actors, as was the case with the Conti group. These actors now use what is called ”big game hunting” (Interview 10). This approach is a planned, long-term operation that targets larger organizations with valuable assets.

For an attack to be successful, the ransomware group must first do a proper reconnaissance of the network environment. The Ransomware actors try to keep under the radar, while they navigate through the network (Interview 10). Ransomware tries to identify the ”crown jewels” of an organization. In this context, the crown jewels refer to the most valuable data and systems. If these data and systems are compromised it would cause big disruptions and thus the organization would be more willing to pay.

The ransom groups try to identify where backups are located within the network but also conduct research on the breached organizations itself (Interview 3). The information gathered includes the organization’s financial standing allows the group to determine a feasible ransom amount. During the identification phase, the ransomware groups try to find out what type of access they have, the specifications of the systems connected to the victim organization’s network, or whether they have access to different file shares. Moreover, they check if these shares have backups (Interview 4). Lastly, they try to identify key network hosts who are responsible for the connection between the regular network and the backup network (Interview 1). To do this, they test the network’s resilience and find out potential vulnerabilities.

After reconnaissance, ransomware actors usually move on to privilege escalation (Interview 3) and lateral movement. The lateral movement here allows the ransomware groups to pivot to other systems (Interview 7) as the ransomware does not only want to gain access but also would like to maintain access over an extended period (Interview 8). They are also noted to target offsite backup storages during their operations (Interview 1)

Once they are inside the network, they often attempt to gain control of a high-level IT admin, backup manager, or related accounts (Interview 1). The escalated privileges can give full control over the system to the groups, especially if they are tied to the active directory (Interview 1). If the active directory is compromised, then the organization will have more troubles regarding the containment and eradication of the ransomware and would need to wipe and rebuild from zero. To get into the high-level admin or backup manager accounts, ransomware groups may utilize keylogging (Interview 1). The keylogging allows the attackers to figure out the passwords of these accounts. These accounts could then be used to delete backups and logs or could lead them to compromise the active directory (Interview 2) (Interview 3). If the active directory is compromised and is inoperable, users will not be able to login or access files which would result in significant operational disruption (Interview 1).

Once the ransomware attackers find the backups and gain access; they can simply infect (poison), encrypt, and corrupt the backups (Interviews 1, 3, 4, 6, 8, 9, 10). They do this by either directly altering the backup files or they could tamper with the management systems that create the backups. Tampering with management systems is especially necessary to attack backups that are usually offline and are harder to erase, such as tapes. Attackers can wipe the tapes through the management utility if they are connected in some way (Interview 3). The attackers are noted to have an easier time attacking tape systems that are automated (Interview 8). They also can make the management system act like nothing is wrong in the system while writing only garbage to the backups to corrupt them (Interview 1). This way, they can wait and stay undetected until they know that the organization will not be able to recover from the backups they have.

It is noted that most of the time, it is more beneficial for a ransomware group to infect the backup files or systems so that once recovery happens, they will still have a way to repeat the attack on the victim (Interview 6,8). This makes it extremely challenging for organizations to be able to also trust their backup data. Therefore, organizations will face increased downtime in order to analyze and check whether the backups may be tainted or not.

Even if the attackers cannot get access and are only partially successful in their prior reconnaissance, privilege escalation, and lateral movement phases, they could use utilities such as the Windows volume shadow copy service or the PowerShell to start wiping partitions that they are aware of, etc. (Interview 3). This creates problems for organizations as the backups will simply not exist for recovery. This information also puts extra pressure on IT Complexity management as then, organizations must proactively secure these utilities and restrict their access from most of their accounts.

4.3. Backup & Recovery Practices

4.3.1. Utilized Frameworks for Backups

Organizations utilize cybersecurity frameworks to guide them with respect to setting up their backup plans and recovery policies. The frameworks are noted as state-of-the-art in the industry. Interview 10 notes that the organizations are recommended to utilize NIST and DORA ”in order to gain insights on what to do in the event of a data breach or what to do if backups are affected and to assess whether that falls into the context of GDPR.” They also note that the most utilized frameworks are the NIST and DORA frameworks, although later reminding that the frameworks utilized will differ per context, sector, and business plans. Interview 9 also notes that the NIST is very popular and states ”Out of 10 organizations, 9 would have based their policies on NIST.”

Which framework had been mentioned by which interview has been given below;

Table 4.1: Mentioned Frameworks

Framework Name	Mentioned In Interview
NIST	1, 2, 4, 6, 8, 9 10
DORA	6, 10
CIS Controls	1, 6
COBIT	2
ISO 27001/27002	2, 4, 6, 8, 9
Third-Party Documentation: Microsoft, Amazon Web Services, Google Cloud Platform	2, 5
DnB Good Practices	6

4.3.2. Backup Strategies

There are many strategies regarding backup and recovery given by the experts.

According to Interview 2 and 9, backups should be stored offsite and they should be offline. The backups that are stored must be made immutable (Interview 3, 6). In support of these, organizations should follow the 3-2-1 strategy (Interview 5, 7). Interview 6 adds that the DORA framework states that full backups must be done minimal yearly, and then after every significant material change. To reduce the data size overhead however, organizations could try to utilize incremental backups (Interview 7). To further reduce overhead a distinction between the criticality of data must be made. This way organizations can focus on only the most critical data (Interviews 7, 8). These critical data could also be supported by real-time data backup systems. Normally however in many organizations who have backup solutions in place the backups are made regularly and automatically usually at midnight (Interview 9). The critical thing here is, to make sure that the backups are verified. With more critical backup data there must be also a higher retention period and a higher number of testing (Interview 2).

During recovery, organizations are better off avoiding system backups and should only try to utilize data backups (Interview 3). This is however contested by Interview 4 as they note that the critical infrastructure (the systems) must also be properly stored in the backups. They argue that the organizations can restart the systems in a dedicated and isolated environment to stay operational for core organizational functions.

Interview 1 notes that many bigger organizations tend to have redundant backup systems which substitute the primary systems in the event of a crisis. Interview 2 in addition advises that the organizations must also have redundant backup storage locations and redundant backup files, especially on the most critical data, to be secured against a major failure.

Interview 10 notes that third party vendors can be outsourced to, in order to ensure that the backup and recovery systems are kept secure. They note the data vaults, and other emerging solutions by vendors such as Dell which provide beneficial solutions for organizations. Interviews 3, 4, 5 and 6, on the other hand argue that cloud backups may also be a feasible cost-effective alternative. The organizations then are able to follow the guidelines set up by these providers to establish their backup strategies. The providers also noted to supply tools and services in order to keep the backups secure. One example is cross-regional backups. Interview 5 explains that cross-regional backups are redundant backups that are placed in availability zones within different regions to increase the resilience for especially sensitive data that requires a large retention period. Although sounding good, it is imperative that organizations ensure that the handling of sensitive data complies with standards.

Interview 6 notes that organizations tend to use all different storage media options to store their backups. They use tapes, hard drives, the cloud or Microsoft Azure environments depending on their organizational goals. Tapes offer a good solution to making the backup data secure as long as the tape management system is monitored and secured properly.

In short, experts note that organizations should focus on establishing criticality metrics, use the 3-2-1 strategy, ensure that there is redundancy in backup data and systems, and that the backups are stored in offsite and offline secure storages.

4.3.3. Create Backup, Recovery & Incident Response Plans

Organizations need strong backup, recovery and incident response plans that are laid out properly way before a crisis unfolds in order to keep downtime to a minimum. However in this regard organizations are noted to need help (Interview 2). The plans will allow an organization to be aware of their backup and recovery capabilities. The plans also make sure that the organization's teams are able to respond in a crisis effectively.

Interviews 1, 2, 4, 5, 6, 7, 8 and 9 note that in these plans one important aspect is to ensure that the criticality metrics for their systems are established in the plans. The most critical system here, noted by the interviewee with the ID 1, is that ”..the active directory, which is the service responsible for access. This system constructs file access, system access, etc. and keeps all the passwords and privileges of the user account”. The plans must highlight all the necessary information such as addresses for the main controllers and Active Directory information with regard to accounts (Interview 6).The other criticality metrics are situational and are dependent on the organizational circumstances.

Interview 4 notes that the critical questions to answer always are ”What systems are important and What needs to be operational?” The classification of the critical data for documentation also allows, organizations to get an understanding into what data they handle. The documentation should also highlight the differences in testing or storing different criticalities of data with regards to backups, also known as the ”tiering” approach. Then, the types of the backups, their frequency and retention rates become major points to consider.

The documented plans must also allocate responsibility, accountability, and operational goals for individuals that will be adhered to during a cyber attack (Interviews 1, 10). In a ransomware incident, there is no ”time to discuss who's responsibility is what" (Interview 10).

Other notes must be made on the recovery time and recovery point objectives in the plans. Moreover, the backup protocols and recovery protocols must be documented very specifically to allow for employees who might have less experience to understand what they must do during a crisis (Interview 4). It is also crucial that the employees are aware and are educated about the backup, recovery and incident response plans. The plans must be clear enough so that different teams who might have different perceptions of backups and recovery are aligned (Interview 9), and thus drive need for communication necessary to a minimum during a crisis. Furthermore, the stakeholder management that must be done during the crisis must be pre-established. The ”Who will be called when?" (Interview 8) question is important for organizations to answer.

Interview 6 also adds the fact that the sequentiality and the dependency of systems or databases must be properly established within the recovery plans. This way, the organizations will know the order in which they must rebuild the infrastructure (Interview 6).

4.3.4. Testing

One pain point in organizations is testing when it comes to backups and recovery. Interview 10 notes that there is a lack in testing and that 50% of the recovery strategy should be about testing and practicing of it. Interview 5 supports this claim as they note that they have experienced inadequate testing in organizations. They further argue that within their experience, organizations tend to be comfortable and say that they have backup strategies and that they comply to frameworks, however, they fail to do dry-runs of the scenario to figure out what actually might happen during a crisis situation.

Interviews 3, 4 and 6 note that proactive testing corresponds to ensuring that the backups are available and valid. A small challenge therein is that the testing might require testing very high volumes of data, which might relate to 4 to 24 hours of recovery process with a 10 gigabytes optic fiber network (Interview 2).

Interview 10 notes that following backup and recovery plan generation, it is crucial that the recovery is tested end-to-end. Without checking if all of the procedures work as planned, organizations would still have gaps in their knowledge. This also relates to having experience, as end-to-end recovery testing allows employees to become more resilient at the time of the crisis, and ensure the execution of the plans during a real crisis. (Interviews 7, 9, 10).

Interview 1 notes that organizations may also want to have regular penetration testing to simulate specific ransomware attacks. As these simulations are based on the goals of the organization, then testing the backups and whether the attackers can access them in order to infect or corrupt them becomes essential.

4.3.5. Maintain Further Cyber Hygiene

Organizations must make sure they set up proper cyber hygiene in order to ensure the security of the backups. The cyber hygiene could be set up by ensuring proper access management and monitoring, consolidating the segmented IT Landscape, and ensuring network segmentation and separation.

One important aspect then is that access management must be set up properly.

Only administrators with high privileges must be able to access systems (Interviews 1,2). The permissions in lower-level accounts must not be overlooked and be monitored to see whether they have permissions to change a higher-level password, etc. (Interview 1). Who can view, alter or transfer the backup data are important questions to always monitor (Interview 2). Organizations must make sure to utilize the zero-trust policy, where all access is severely limited to accounts and only if it is necessary, access is authorized. A new strategy that could be taken with access is that if access is granted for example outside of business hours it could be flagged as an indicator of compromise (Interview 8).

The segmented IT landscape must be consolidated (Interview 10). Organizations try to now trim down the number of applications, and third-party suppliers, necessary for their operational continuity and they try to utilize integrated tools. These integrated tools should be acquired from vendors who the organization already is satisfied with.

Network segmentation and separation are key measures that organizations tend to establish and if they don't, they must, which results in different tiers of networks and subnets (Interviews 2,4,6,7 and 8). The backups must always be separated from the primary domain. The network segments must have differing policies, and organizations could introduce group policy plans where users on a specific segment of the network are only barred to do certain operations and actions, and they cannot use third-party libraries (Interview 10). Again the zero-trust policy and always assuming the eventual compromise of accounts are important paradigms that organizations need to keep in mind.

4.4. Backup & Recovery Challenges

4.4.1. Situational & Complexity

There are many challenges that organizations face during backups and recovery. The first idea is to understand that the field is very situational and complex with a lot of stakeholders involving a lot of middle management. Although the basics are similar, each solution is unique to the organization and many solution mixes could be made ransomware-proof. The variability in technological and organizational choices, and the constant trade-offs make it hard for organizations to navigate the field (Interviews 2, 3, 5, 6, and 8).

Interview 1 states that it is always difficult to restore systems from zero and the challenges each organization faces will be different with regards to their backup configurations. However, the complexity, technology and different services will make it time-consuming to rebuild whilst still having the need to do the analysis of the attack (Interview 6). The downtime of the analysis of the attack is also dependent on how big the network logs are, and their retention period (Interview 7). This is because the sheer size of the logs might add up to what must be analyzed.

Related to the situational and complexity of the domain, it is also noted that infrastructure and resource management is a big hassle. One of the reasons being is that there is insufficient capabilities and resources (Interviews 1,2,9). Interview 1 states that many organizations prefer to utilize the cost-effective cloud solutions as in-house development usually costs a lot of time and resources. This is supported by Interview 7 as they note that in-house response teams are simply too expensive.

4.4.2. Infrastructure & Resources

Interview 1 notes as the main issue that there are not enough resources or the allocation of resources to establish strong backup and recovery systems. According to Interview 2, the overhead of IT is too large for organizations to invest in. They are too focused on their core value offering that ”they do not have the time to push for backups and recovery”. Interview 4 notes that the hardware [and software] should be maintained and regularly updated. This management also creates considerable IT Overhead and result in both capital and labor costs.

There is also a lot of dependencies in the environment. Interview 1 notes that if there is an old version dependency after an update for a third-party library or a system component, the necessary data might not be in the right spot, or usable by the new update. This might cause organizations to have to waste resources in order to ensure compatability.

The dependencies however are not necessarily only about version control. It is also about third parties. Interview 2 notes that if the cost of setting up the IT team, in-house backup, and recovery practices outweigh the cost of paying the third party, then the organizations will just outsource it. However, when cybersecurity practices such as backups and recovery are subcontracted, there are concerns about whether the third party will adhere to security or regulatory requirements over their data, as they do not have control or insight over what happens to it anymore (Interview 2). These third parties may also be slow in response during a crisis scenario. Interview 3 states ”[organizations] will need to hope that they (the third-party) do not get breached at the same time.”

Another big aspect is IT Integration, monitoring, and security. The IT Landscape has just gotten too big, which creates a segmented IT landscape (Interview 10). With this, organizations must watch out for many different applications from perhaps many different vendors and see how they can integrate well with one another. To handle large loads, however, organizations need many different applications (Interview 6). Although if they start intertwining during a recovery scenario, then this is a problem. The IT Landscape is also not only restricted to the organization’s applications. There is a lot of shadow IT (Interview 4) in the environment.

The IT Landscape then creates vulnerabilities in security. One way to address this is to ensure proper monitoring. Interview 3 notes that IT systems must be monitored by humans. This is the case even when there are a lot of defense systems in place. If the IT employees do not monitor alerts, then the defenses will not be effective (Interview 3). The IT employees must also monitor unpatched software and other vulnerabilities that might be prevalent in the network. This monitoring also refers to ensuring the usability of backups by verifying them (Interview 8).

Another issue is the fact that during recovery many events are time-restricted (Interview 10). For example, the IT equipment management to ensure that the equipment is cleaned and analyzed after an attack would take time (Interview 1). There are also a lot of analyses that must be conducted. Interview 4 notes that the organization must find out by asking the right questions and conducting analysis to ensure that the attack is not an insider threat, that the ransomware attackers do not have access anymore, etc. Interview 7 reminds that the analysis of the attack is one of the most critical moments for potential future litigations. Interview 6 notes that even though analysis takes time, the rebuilding of the complete infrastructure will also take time. The duration will be related to the time to spin up new machines, etc.

Interview 2 also noted that some organizations are immature regarding cybersecurity and that they do not have network segmentation or separation. This then contributes to ransomware being able to move laterally while escalating their privileges.

4.4.3. Organizational Orientation

Interview 1 notes that organizations tend to see IT as a cost and not as a win. If organizations are not hit by a ransomware attack, then they will not be benefiting from their investment. They also note that many critical problems for cybersecurity are not fixed for prolonged periods in organizations. The organizations prefer to use temporary solutions, and ”temporary solutions tend to be always made permanent.” (Interview 1) This, with the addition of the cybersecurity insurances against ransomware (Interview 1) would make the organization less likely to invest in cybersecurity maturity.

Businesses in general try to move things forward (Interview 5) and focus more on their core value offering. This makes them less likely to take a step back and establish plans to train the employees and set up processes for backup & recovery (Interview 5). This orientation could very well in fact make them prone to only being reactive to cybersecurity incidents. Many organizations indeed are quite reactive and only think about cybersecurity when crisis strikes (Interview 3).

Organizations will also try to reduce costs by outsourcing backups to third parties which could generate the third-party risks previously mentioned. Interview 2 supports this claim as from their experience they note that an organization would choose to pay to outsource then create in-house IT teams capable of backup & recovery operations, being especially true if the organization is not IT-oriented.

4.4.4. Coordination & Communication Issues

Organizations also have trouble with coordination and communication. Mainly organizations lack in some way or form Incident Response, Backup and Recovery plans. These plans entail characteristics of the backups and the recovery processes. Without a well-established plan that defines the criticality, sequentiality, and inter-dependencies of backup data and systems, and their prioritization during recovery; an organization will be confused during crisis.

Interview 1 notes that a bottleneck in recovery is not having recovery process guidelines. Interview 2 states that ”many organizations still do not have very formalized and well-defined backup strategies and policies.” Interview 3 supports this by stating that though organizations do have backups, many of them do not know their recovery processes. Interviews 5, 7, 8, 9 also approve this claim. Many organizations also do not know what their critical systems are (Interview 3). Interview 4 notes that the critical and the correct dependencies of the systems to be rebuilt must be properly established and mapped.

Interview 1 states that they know from experience that many managers lack knowledge regarding recovery. This is supported by the claim from Interview 2, as they note that there are a lack of professionals who are experienced in backup and recovery.

4.4.5. Backup Related

Organizations also face challenges when accessing their backups during a crisis. This is due to either a loss in backup integrity or is related to the fact that the backup data is inaccessible. Interview 2 notes that many organizations have their backups next to the primary production domain which makes it easier for attackers to find and access (Interview 3), and also relates to the fact that they become inaccessible when the main systems are down. Moreover, if an attacker gets access to backups; they might get sensitive information including the admin credentials, directories, etc. in the network. This way they can use the backups to further target either redundant copies, or create more intricate tactics such as executing the payload on certain types of runtime, etc. (Interview 1)

For the challenges relating to backups, there is also the fact that there is simply too much data to be saved in the backups and to be recovered. The data size is a huge pain point as organizational capabilities cannot match it (Interview 2). This makes it so that then, organizations must discuss how much data loss is tolerable, and perhaps utilize methods of data reduction.

4.4.6. Insufficient Testing

It is also evident from the interviews that organizations do not do enough testing (Interviews 1,2,3,4,8 and 9) of their backups & recovery. The testing is stated to be done ad hoc or is minimal (Interview 2). It is noted that many organizations do their backups and hope they work (Interview 3). Interview 9 also adds that it is not only the recovery of the backups but that there are gaps in organizations regarding end-to-end recovery testing in order to see whether all recovery processes are set up correctly and whether the organization can achieve the RTO and RPOs. These then make it so that organizations do not know if their backups will work, or whether they will be able to orchestrate a strong ransomware attack response and recovery.

4.5. Conclusion

This chapter has gone over the findings generated through 10 semi-structured expert interviews. The experts have been found through networking, and contacting through LinkedIn and email. The findings have been classified in three focus points which highlight each of the first three subquestions.

The findings regarding the first subquestion, ”How do the most salient ransomware families target and compromise backups?”, in this chapter is that ransomware currently has evolved into ”big game hunting” by organized crime groups, also called ATPs. These are planned and long term operations targeting larger organizations. They do reconnaissance of the network environment to be able to have an insight into the crown jewels of an organization. They then keep under the radar while doing lateral movement and privilege escalation to achieve accounts with high permissions. When they have the access level required they find the backups to infect (poison), encrypt, or corrupt them. They also utilize legitimate operating system tools to destroy shadow copies or wipe other partitions that hold backup data.

The findings regarding the second subquestion, ”What are the current state-of-the-art backup & recovery practices that are (or can be) employed by companies to safeguard their data and ensure operational continuity?” are that: firstly, it is important to note that the most used identified cybersecurity frameworks in the industry are NIST, DORA, CIS Controls, COBIT, ISO27k Family, Third-party documentation by Microsoft, Amazon Web Services, Google Cloud Platform, and lastly DnB Good Practices.

On the backup strategies side; backups must be stored offsite and offline, they must also be made immutable with respect to organizational goals. Backups must be made redundant, other copies must be available should one backup be comprimized. Organizations should follow the 3-2-1 strategy, and full backups of the data must be done minimal yearly. Ensuring that the criticality, the sequentiality, and the dependencies of the data are known and well documented will allow organizations to not face additional downtimes or commit to mistakes during recovery. The backups should be further protected by utilizing solutions such as data vaults, cloud backups (offsite) which can offer cross-regional backup capabilities. The organizations also can pick from many backup storage options with respect to their organizational goals. What is essential however is that backup, recovery & incident response plans must be well thought out and documented. Responsibility, accountability and the establishment of operational goals to be adhered to, during a crisis moment is crucial. Organizations must also always test their backup and recovery capabilities end-to-end, starting from the verification of the backup to a full-scale organizational test to make sure that if a real crisis strikes the organization can operate under stress and function as a well-oiled machine. Further cyber hygiene should also be set up through proper access management, consolidating segmented IT landscape, and ensuring proper network segmentation and seperation. Here, the zero-trust policy is essential to comply to in all levels of IT in the organization.

The findings found under this chapter regarding the third subquestion, ”What are the main challenges faced by organizations regarding backups & recovery?” are that: the backup & recovery issue is quite situational and complex. The issue also requires well laid out IT infrastructure & high amounts of resources, there are also organizational orientations that are incompatible with cyber resilience regarding backups & recovery. Moreover, there are coordination & communication issues between IT teams, issues when creating or accessing backups, and lastly challenges faced due to insufficient testing of backups and end-to-end organizational recovery.

5
Synthesis & Discussion

This chapter maps out which backup strategies stop which ransomware attacks and makes recommendations on the challenges that the organizations have been found to face. The reflection is important to finally answer the fourth subquestion, namely, ”What are the implications of the discovered findings with respect to providing a recommendation list to organizations?” which is the precursor in order to answer the main research question by inducing a reflection of all the findings together to create one final list of recommendations. The reflection on all the findings together is made under Section 5.1. Then following this, a final list of recommendations for organizations is given under Section 5.2 in order to answer the main research question ”What recommendations can be made for organizations to help them develop good backup and recovery strategies against the growing ransomware threat through a systematic and methodological approach?”.

5.1. Implications for Recommendations

It is important to note that many organizations face challenges with respect to the lacking of the associated backup strategy. This is like a mirror. For example, inadequate access management which is a major pain point in organizations, could be solved by doing proper access management. The lack of testing, which is another challenge, could be fixed by doing verification of the backups after they are created and scheduling recovery tests of the backups. The testing and verification will also ensure that organizations are aware of whether they have incorrect configurations of backups. The lack of up-to-date backups could be solved by making the recovery point objectives smaller and perhaps by utilizing backup solutions such as BaaS that will allow for automatic backups and the automatic verification of the backups. The organizations could also utilize SaaS solutions and their C2C environments to automatically backup into the corresponding vendor’s cloud.

The backup & recovery practices had been categorized under backup storage options and operational practices & procedures. The backup storage options here, that are chosen will be based upon organizational plans that are very dependent on the situation. Many mixes of backup storage options can be made ransomware-proof. Therefore the choice of the storage only reflects the decision-making on the costs, their pros and cons, and limitations. Further insight is given about the data sizes, transer speeds and compatibility for these storage options. Organizations do face challenges with regard to finding sufficient capacity to continue their backup operations. Hereby, the choice must create less complexity, should be easy to secure, and could be made available during a crisis quickly. The limitations of the backup storage could be fixed in a case-by-case scenario. For example for the cloud, a pain point is that due to the pay-as-you-go business model, costs can add up. Here shared storage in a public cloud can be utilized. However, then this relates to privacy concerns. It could be seen that there are many tradeoffs and a single right strategy is not possible. However, organizations must ensure the proper confidentiality, integrity, and accessibility of their backups.

On the other hand, the operational practices & procedures are the main categories of; protection, generation, backup & recovery policy, testing recovery & employee training, Utilizing SaaS and BaaS, Data Reduction, Recovery Principles, and Collaboration. Protection through access and privilege management is directly related to the mitigation of reconnaissance, lateral movement, and privilege escalation processes of the ransomware attacks. If correct access management is done through implementing zero-trust strategies such as giving out minimum access rights and privileges based on roles, establishing multifactor authentication and quorums; the ransomware actors will have a very hard time getting access into high-privilege accounts and waste their time in the network which will give more time for detection to happen. Furthermore, if all roles are given specific operational controls, and that they cannot for example, enumerate the networks through their accounts ransomware; they will be less able to use accounts to move laterally in the network. Here one interesting insight could be that if monitoring comes into play regarding access, if any action towards an account's credentials or a non-normal operation from an user account such as connecting externally after work hours are flagged as potentially malicious; then a SIEM or a SOC can pick up on these and identify them as indicators of compromise. The non-normal operations could also be extended into not allowing any user but only key admins to terminate the database, backup solutions or antivirus softwares.

Moreover, the reconnaissance, lateral movement and privilege escalation of ransomware could be stopped if organizations keep good cyber hygiene, and segment and separate their networks. Segmentation will allow to control the traffic flows to be controlled between different subnetworks through having different security policies and firewalls. Network separation through logical or physical means of the backups are crucial as then there will be no direct communication paths between the primary domain and the subnetwork that contains the backup data. The primary domain must never have any credentials or access to the backup subnetwork under any circumstance. Only in the event of a crisis should the separation be lifted as necessary for recovery. The monitoring aspect also comes into play here as they must look into the network traffic flows to identify any indicators of compromize. This could be done for example by checking the network logs and noting the connections made to or by external sources which are not part of the secure connections list of that particular network segment. If these are turned into alerts and that SIEM or SOCs can note them, then ransomware will be contained prior to them reaching to the backups.

Organizations must ensure proper patch management. Patch management is directly related to hampering the process of using vulnerabilities by ransomware. Furthermore, organizations must ensure to disable all unnecessary services and protocols such as the vssadmin, wmic, etc. that the ransomware attackers can utilize to jeopardize the system and the backups.

Should backups be accessible through a high-level access account by the ransomware; then they can either corrupt, infect or delete the backup files. However, they are also noted to exfiltrate the backup data, which might have very sensitive information including the access credentials or other information in the network of the organization. To circumvent this encryption is possible, however, encrypting the backups will not allow organizations to be safe against their backups being destroyed in recovery because the ransomware attackers can utilize secondary encryption. To stop deletion of the files here a protection mechanism could be to utilize a backup solution or a backup storage that allows for a soft deletion. Then the files will remain in a cache which could be recovered.

Isolating recovery data through air-gapped offline, offsite solutions, and utilizing redundant, immutable backups will jeopardize the ransomware attacks as an air-gapped copy, which is either physically (offline) or logically separated from the main system could not be accessed by the ransomware and could still be utilized for recovery. Immutable backups also protect against ransomware tampering with the backup integrity. Redundancy on the other hand, ensures that even if a backup storage system is compromised there is still another. This is the method utilized in both popular 3-2-1 and the unique 4-3-2-1 strategies.

Regarding recommendations for the challenges, it could be seen that organizations have a lot of problems with complexity management. There could be huge IT Overheads created due to; having multiple backup solutions, where some are legacy, or having complex data virtualization environments, or managing the compatibility and the tiering of storage. Furthermore the IT Overheads are caused by huge segmented IT Landscapes. Some strategies to combat this challenge are that first, organizations must also manage the patching on their third-party software to ensure their complex IT Landscapes do not relate to a expanded attack surface where the vulnerabilities might be exploited by ransomware. They must also ensure compatibility of the previous versions of the backups that still needs to be retained, with the up-to-date versions of the backup solution. Organizations should target to streamline their IT Landscapes and consolidate functions under integrated applications that vendors provide.

There are some challenges that are created by the perceptions and the approach the cybersecurity by organizations. For example, organizations have been found out to sometimes not take action on critical cybersecurity issues. Organizations are also prone to applying reactive and temporary fixes, noting that cybersecurity is a problem for later but then making the temporary fix a permanent fix. Organizations also are more willing to pay then create in-house, hybrid or offsite environments to sustain their backups. They note that the attention taken away from their core value offering is higher in marginal costs than actually creating a strong backup & recovery infrastructure. These are the realities mostly caused by the ”cybersecurity is a cost” thinking. It is not every organization who do this, but the organizations that think this way will have gaps in their backup & recovery practices. The way to fix this is developing by itself. It is noted by Interview 8 that experts are seeing a change in organizational thinking. Therefore, challenges are there but it is getting better. Organizations must change their thinking to be proactive regarding cybersecurity. Sometimes however, it comes also down to not simply not being aware. In the Maastricht University case, FOX-IT has noted that [170] the university planned to map their devices in the network.. after the attack, which introduces additional downtime. Backup and recovery policies, procedures that include all information such as this must be documented prior to attacks.

Regulatory and Third-party provider challenges could be managed by ensuring that the third parties being utilized for their backup storage or applications comply with the specific regulations. This could be done by doing assessments of the provider's business model and monitoring the contractual obligations. The responsibilities of the third parties and the organization must be laid out explicitly and also need to be part of the documented backup & recovery policies. This way, organizations can choose third parties who comply with their processes and regulations. Furthermore, this monitoring should not end after the contract is finalized but must be revised with respect to any changes that happen.

Organizations have trouble ensuring proper coordination and communication during crisis. This also relates to the lack of knowledge and lack of cybersecurity awareness and training. These issues could be solved primarily by establishing and documenting backup & recovery plans and policies that are also used in training the employees. With this in place, employees will understand the procedures that they must undertake, establish responsibility and accountabilities, more importantly gain insight to what is the prioritization of recovery, and what are the dependencies, sequentiality or write-order fidelities for a proper rebuilding of the infrastructure. End-to-end recovery tests of these plans and policies then also relates to having a strong stance against a ransomware attack. Moreover, organizations could invest in crisis event management software which is offered by vendors such as Blackberry to ensure that communication stays intact in the case of a crisis. These vendors now have updated their value offerings to contain cyber attack incidents.

5.2. List of Recommendations for Organizations

After the synthesis of ransomware attacks and challenges that organizations face regarding backup & recovery, with their corresponding practices through a reflection that has answered the last subquestion in the previous section; the main recommendation list is given in Table 5.1.

Table 5.1: Final Recommendations for Organizations

Category	**Subcategory**	**Properties**
Generation		
	Isolate Recovery Data	
		Use Offline, Cloud or Offsite Systems and Services
		Utilize Airgapping Methods
		Manage and Periodically Assess Offsite Arrangements

Table 5.1: Final Recommendations for Organizations

Category	Subcategory	Properties
	Do Regular Backups	
	Use the 3-2-1 Strategy	
	Create Immutable Backups	
	Ensure Compatability Correct Configuration	
	Establish Failover Backup Systems	
	Create Redundant Backups	
	Use the 4-3-2-1 Strategy	
	Continuous Data Protection (CDP) Backups	
	Ensure at least one Full Backup	
Protection		
	Access and Privilege management	
		Implement MFA
		Only Give Minimum Access Rights and Privileges
		Utilize a Zero Trust Access Strategy
		Create Quorums
		Physically Secure Devices and Media
		Seperate Roles and their Duties
		Disable Plaintext Passwords
		Enable Credential Guard
		Encrypt Login Credentials at Rest
		Instead of passwords use SSH
		Establish Proper Authentication
		Establish Role-based Access Controls
	Encrypt Backups	
		Establish Key Management
	Monitor Domains, Backups Recovery Data and Application Health	

Table 5.1: Final Recommendations for Organizations

Category	Subcategory	Properties
		Create Network and Data Plane Segmentation and Seperation
		Proper Patch Management
		Disable all Unnecessary Services and Protocols
		Streamline IT Landscape
		Utilize Soft Delete
	Testing Recovery & Employee Training	
	Backup and Recovery Policy	
		Content, Criticality, Prioritization and Characteristics of Backups
		Establish Operational Plans Regarding Saving Data
		Do both Internal and External Periodic Audits
		Establish Responsibility
		Keep an inventory of Backup Storage
		Establish a retention period
		Establish Communication and Cooperation Plan
		Establish KPIs and Requirements
		Backup Lifecycle Management
		Consider Tiered Approach for Backups
	Utilize SaaS and BaaS, or other utilities from third-parties to reduce costs and overhead	
	Data Reduction	
		Utilize Data Deduplication
		Utilize Data Compression
	Utilize Recovery Principles	
		Use a Version of Backup that Predates Infection
		Wipe Systems and Rebuild From Good Backups
	Collaborate with Other Organizations	

Table 5.1: Final Recommendations for Organizations

Category	Subcategory	Properties
		Consider Using a Crisis Event Management Tools
		Establish SIEM or SOCs for Better Monitoring
		Move from ”Cybersecurity as a Cost” thinking to a proactive ”Cybersecurity is a Must” paradigm
		Governments and other authoritary entities such as the EU should provide incentives for strong cybersecurity adoption.

The main recommendation list therefore is the conglomeration of 203 documents that mention the ransomware attacks towards backups, best backup & recovery practices, and challenges that organization face. The total of 203 documents have been extracted from; cybersecurity frameworks, antivirus and backup software provider reports, cybersecurity blogs, academic literature, and expert interviews. The recommendation list finally answers the main question ”What recommendations can be made for organizations to help them develop good backup and recovery strategies against the growing ransomware threat through a systematic and methodological approach?” The list of recommendations above has been generated through a rigorous, systematic and methodological approach consisting of the three methods.

6 Conclusion, Limitations & Future Research

6.1. Conclusion

The thesis has given organizations a set of recommendations that is supported by a comprehensive number of materials in the domain. To this end, four types of knowledge sources have been utilized. These sources were; academic literature, established cybersecurity frameworks, reports of antivirus and backup provider organizations, and interviews. Correspondingly, the thesis has utilized three methods; semi-systematic literature review, qualitative content analysis, and semi-structured interviews in order to do a qualitative analysis of the data acquired from the sources.

The main findings in order to answer the first subquestion of the thesis, namely, *”How do the most salient ransomware families target and compromise backups?”* were that ransomware target backups during their reconnaissance, lateral movement and privilege escalation processes. They attempt to get access by utilizing vulnerabilities and tools to be able to alter backup data so that the organization does not have a recovery option available. Once they do have access; they corrupt, delete, encrypt or infect the backup data.

To answer the second subquestion, namely, *What are the current state-of-the-art backup & recovery practices that are (or can be) employed by companies to safeguard their data and ensure operational continuity?* organizations are then advised to stop these attacks as much as possible by making a mix of solutions that follow state-of-the-art backup strategies to allow for redundant, immutable and air-gapped backups in offsite and offline storages. Moreover organizations should sustain proper cyber hygiene by; ensuring correct access and privilege management, patch management, monitoring networks, segmenting and separating networks, and encrypting the backups. Furthermore, organizations must ensure all these backup & recovery processes and policies are documented explicitly and properly. The testing of these process and policies, including the testing of backups must be made mandatory.

The main challenges that organizations face during backups and recovery, which answers the third subquestion of the thesis, asking *”What are the main challenges faced by organizations regarding backups & recovery?”* are related to; complexity management, segmented IT Landscape, regulatory and third-party provider challenges, inadequate security and backup policies, high costs and allocation of resources, and challenges of the backup solutions. It is also noted that organizations may be overconfident and trust in their backups more then they should. They also have been found out to not fix many crucial pain points for their backup & recovery practices due to their focus only being on their core value offering.

The findings have been translated into recommendations with the answering of the fourth subquestion, *”What are the implications of the discovered findings with respect to providing a recommendation list to organizations?”*, as the final piece of the puzzle prior to completion in Chapter 5. Thus with this, the thesis has generated a recommendation list through a systematic approach which answers the main research question, *”What recommendations can be made for organizations to help them develop good backup and recovery strategies against the growing ransomware threat through a systematic and methodological approach?”*. The recommendation list is a main deliverable of the thesis for organizations in both private and public sectors to be aware of, and to have insights into what they need to ensure that they do not face many of the same challenges and attacks by ransomware that other organizations face.

6.2. Validity

Validity considerations on the research have been made on four criteria of assesment noted by [30]. These are: credibility, transferability, dependability and confirmability. Through triangulation of the literature review, the content analysis and the interviews the credibility and confirmability of the research has been established. Credibility refers to internal validity, which in this context, looks at whether the important factors have been identified with respect to answering the research questions. Considering how similar things have been highlighted per method to answer the same subquestions throughout the thesis, credibility of these findings is well established.

The confirmability aspects relates, in qualitative content analysis, to determine if the data supports the conclusions. It is related to see whether there is conceptual consistency in the thesis. The implications and the conclusions per chapter has been conceptually consistent with the findings as they are the ground on which the recommendations are built upon. The data thus supports the conclusions.

Transferability on the other hand is established by reaching data saturation through a variety of sources. This has been achieved through identifying 82 sources in the literature review, 102 sources in qualitative content analysis, and 10 expert interviews. The research is rigorous enough, and has been comprehensive enough to start generating duplicate findings.

Lastly, dependability is the most challenging due to the replicability of the study. As described before, a qualitative research is very dependant on its researcher. This indeed does create researcher bias with regards to the coding and the categorization of these codes. However, considering the amount of times that certain concepts were mentioned in similar sentences, it could be seen that in a lot of literature the ways the concepts are understood are similar. Therefore, the research can indeed be mostly replicated with similar findings and implications considering that the same systematic and methodological approach is used. It is however a possibility that another researcher can come up with their own categorization and abstractions of the findings.

6.3. Limitations

The thesis has had certain limitations. Each limitation is given with their mitigation (if applicable) next to them below. The limitations are:

- Semi-Systematic Literature Reviews have focused on articles specifically about ransomware, backups, and challenges. There could have been more research on disaster recovery, as it is an overarching concept for cyber recovery.
- The categorization of the concepts was quite difficult as most are interrelated concepts and due to nuances many cannot be merged into each other or made mutually exclusive.
- The generated open coding has not been benchmarked with respect to inter-coder reliability.
- The qualitative content analysis method is specifically tied to the skills of the researcher. As the researcher is a student of the content as well, the coding confidence, etc. might not have been uniform.
- Purposive sampling in the qualitative research phase might have generated researcher bias.
- The number of interviews and the sampling of the interviewees could be expanded in order to get more industry insights.
- Material that could not be scraped from the internet by using Nvivo Ncapture, or downloaded were excluded from the study.
- Some material that was clearly biased and promotional products were also excluded from the study. The reason for this is although it is true that many backup solution providers and antivirus companies have highlighted attacks or challenges that their solution directly fixes, some have gone quite overboard were mixed a lot of opinions.
- The Cybersecurity Frameworks also have their corresponding implementations, which may have been missed during their analysis. This could be for example, article 12 in DORA regulations state the necessity of backups and their protection. However, the specific implementation of this needs to be established through the understanding and experience of implementation in other directives like the NIS2 or DnB protocols, and understanding what the EU supervisory board asks.

6.4. Future Research

The thesis highlights some avenues that future researchers can follow. Future research can be done by:

- Choosing a category and doing a more in-depth analysis of it. Each category could actually be made quite complex. Focus could be made on the technicalities, implementation, or limitations of that specific category, on different industries etc.
- Using the generated codebook to analyze other cybersecurity frameworks or other related or additional material in a deductive fashion to identify trends or discourse in the future. Here, establishing inter-coder reliability can be a main concern to benchmark the replicability and external validity of the coding.

- Each cybersecurity framework, especially the DORA regulations(which is mandatory starting 2025), could be operationalized in the context of backups in a case study.
- Interviews can be expanded and classified with respect to the perceptions of IT, database management or cybersecurity professionals of different experience levels, and sectors in order to identify differences or similarities in thinking that contribute to challenges that organizations face with regards to backup & recovery.

References

[1] IBM Security, *Ibm security x-force threat intelligence index 2023*, 2023. [Online]. Available: `https://www.ibm.com/reports/threat-intelligence`.

[2] R. Southwick, *Fbi: Healthcare hit with most ransomware attacks of any critical sector*, Mar. 2023. [Online]. Available: `https://www.chiefhealthcareexecutive.com/view/fbi-healthcare-hit-with-most-ransomware-attacks-of-any-critical-sector`.

[3] Cybersecurity & Infrastructure Security Agency, *Ransomware 101 | cisa*. [Online]. Available: `https://www.cisa.gov/stopransomware/ransomware-101`.

[4] IBM, *What is ransomware?* 2022. [Online]. Available: `https://www.ibm.com/topics/ransomware`.

[5] J. A. Abraham and S. M. George, “A survey on preventing crypto ransomware using machine learning,” *2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT)*, Jul. 2019. DOI: `https://doi.org/10.1109/icicict46008.2019.8993137`.

[6] A. Alzahrani, A. Alshehri, H. Alshahrani, and H. Fu, *Ransomware in windows and android platforms*, 2020. DOI: `10.48550/ARXIV.2005.05571`. [Online]. Available: `https://arxiv.org/abs/2005.05571`.

[7] A. Alqahtani and F. T. Sheldon, “A survey of crypto ransomware attack detection methodologies: An evolving outlook,” *Sensors*, vol. 22, no. 5, p. 1837, Feb. 2022. DOI: `10.3390/s22051837`. [Online]. Available: `https://doi.org/10.3390/s22051837`.

[8] N. M. Chayal, A. Saxena, and R. Khan, “A review on spreading and forensics analysis of windows-based ransomware,” *Annals of Data Science*, Jun. 2022. DOI: `10.1007/s40745-022-00417-5`. [Online]. Available: `https://doi.org/10.1007/s40745-022-00417-5`.

[9] B. A. S. Al-rimy, M. A. Maarof, and S. Z. M. Shaid, “Ransomware threat success factors, taxonomy, and countermeasures: A survey and research directions,” *Computers & Security*, vol. 74, pp. 144–166, May 2018. DOI: `10.1016/j.cose.2018.01.001`. [Online]. Available: `https://doi.org/10.1016/j.cose.2018.01.001`.

[10] H. Oz, A. Aris, A. Levi, and A. S. Uluagac, “A survey on ransomware: Evolution, taxonomy, and defense solutions,” *ACM Computing Surveys*, vol. 54, no. 11s, pp. 1–37, Jan. 2022. DOI: `10.1145/3514229`. [Online]. Available: `https://doi.org/10.1145/3514229`.

[11] N. Aldaraani and Z. Begum, “Understanding the impact of ransomware: A survey on its evolution, mitigation and prevention techniques,” in *2018 21st Saudi Computer Society National Computer Conference (NCC)*, IEEE, Apr. 2018. DOI: 10.1109/ncg.2018.8593029. [Online]. Available: https://doi.org/10.1109/ncg.2018.8593029.

[12] G. McDonald, P. Papadopoulos, N. Pitropakis, J. Ahmad, and W. J. Buchanan, “Ransomware: Analysing the impact on windows active directory domain services,” 2022. DOI: 10.48550/ARXIV.2202.03276. [Online]. Available: https://arxiv.org/abs/2202.03276.

[13] CERTNZ, *How ransomware happens and how to stop it*. [Online]. Available: https://www.cert.govt.nz/it-specialists/guides/how-ransomware-happens-and-how-to-stop-it/.

[14] A. Din, *Some of the companies affected by ransomware in 2021*, Oct. 2021. [Online]. Available: https://heimdalsecurity.com/blog/companies-affected-by-ransomware/.

[15] R. Moody, *Map of worldwide ransomware attacks (updated daily)*, Mar. 2022. [Online]. Available: https://www.comparitech.com/blog/information-security/global-ransomware-attacks/.

[16] D. Braue, *Global ransomware damage costs predicted to exceed $265 billion by 2031*, Jun. 2021. [Online]. Available: https://cybersecurityventures.com/global-ransomware-damage-costs-predicted-to-reach-250-billion-usd-by-2031/.

[17] Verizon, *2022 data breach investigations report*, 2022. [Online]. Available: https://www.verizon.com/business/resources/reports/dbir/.

[18] CDW, *Ransomware attacks in the energy industry*, May 2022. [Online]. Available: https://www.cdw.com/content/cdw/en/articles/security/ransomware-attacks-energy-industry.html.

[19] E. Berrueta, D. Morato, E. Magana, and M. Izal, “A survey on detection techniques for cryptographic ransomware,” *IEEE Access*, vol. 7, pp. 144 925–144 944, 2019. DOI: 10.1109/access.2019.2945839. [Online]. Available: https://doi.org/10.1109/access.2019.2945839.

[20] B. Posey, J. A. Miller, and G. Kranz, *What is backup (data backup)?* Oct. 2021. [Online]. Available: https://www.techtarget.com/searchdatabackup/definition/backup.

[21] A. Adshead, *Almost all ransomware attacks target backups, says veeam | computer weekly*, May 2023. [Online]. Available: https://www.computerweekly.com/news/366538492/Almost-all-ransomware-attacks-target-backups-says-Veeam#:~:text=Data%5C%20stored%5C%20in%5C%20backups%5C%20is.

[22] Y.-S. Chen, J.-L. Chou, Y.-S. Lin, Y.-H. Hung, and X.-H. Chen, “Identification of smes in the critical factors of an is backup system using a three-stage advanced hybrid mdm–ahp model,” *Sustainability (Switzerland)*, vol. 15, no. 4, 2023, cited By 0. DOI: 10.3390/su15043516. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85149239564&doi=10.3390%5C%2fsu15043516&partnerID=40&md5=eb38e265ca9ddfcf69911a71466aa28c.

[23] E. A. Al-Qarni, “Cybersecurity in healthcare: A review of recent attacks and mitigation strategies,” *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 5, 2023. DOI: 10.14569/ijacsa.2023.0140513. [Online]. Available: https://doi.org/10.14569/ijacsa.2023.0140513.

[24] H. Neprash *et al.*, “Trends in ransomware attacks on us hospitals, clinics, and other health care delivery organizations, 2016-2021,” *JAMA Health Forum*, vol. 3, no. 12, E224873, 2022, cited By 1. DOI: 10.1001/jamahealthforum.2022.4873. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85145609694&doi=10.1001%5C%2fjamahealthforum.2022.4873&partnerID=40&md5=5612080f121eda53c8328367596b20bf.

[25] L. Ifflander, A. Dmitrienko, C. Hagen, M. Jobst, and S. Kounev, *Hands off my database: Ransomware detection in databases through dynamic analysis of query sequences*, 2019. DOI: 10.48550/ARXIV.1907.06775. [Online]. Available: https://arxiv.org/abs/1907.06775.

[26] V. Kaushik and C. A. Walsh, “Pragmatism as a research paradigm and its implications for social work research,” *Social Sciences*, vol. 8, no. 9, p. 255, Sep. 2019. DOI: 10.3390/socsci8090255. [Online]. Available: https://doi.org/10.3390/socsci8090255.

[27] S. Elo and H. Kyngäs, “The qualitative content analysis process,” *Journal of Advanced Nursing*, vol. 62, no. 1, pp. 107–115, Apr. 2008. DOI: 10.1111/j.1365-2648.2007.04569.x. [Online]. Available: https://doi.org/10.1111/j.1365-2648.2007.04569.x.

[28] U. Pfeil and P. Zaphiris, “Applying qualitative content analysis to study online support communities,” *Universal Access in the Information Society*, vol. 9, no. 1, pp. 1–16, Aug. 2009. DOI: 10.1007/s10209-009-0154-3. [Online]. Available: https://doi.org/10.1007/s10209-009-0154-3.

[29] J. W. Drisko and T. Maschi, *Content analysis*. New York: Oxford University Press, 2016, ISBN: 9780190215491.

[30] M. D. White and E. E. Marsh, “Content analysis: A flexible methodology,” *Library Trends*, vol. 55, no. 1, pp. 22–45, 2006. DOI: 10.1353/lib.2006.0053. [Online]. Available: https://doi.org/10.1353/lib.2006.0053.

[31] B. Enserink, *Policy Analysis of Multi-actor Systems*. TU Delft, 2009.

[32] M. J. Page *et al.*, “The prisma 2020 statement: An updated guideline for reporting systematic reviews,” *British Medical Journal*, vol. 372, no. 71, Mar. 2021. DOI: https://doi.org/10.1136/bmj.n71.

[33] M. Ahmad and W. Elmedany, "A review on methods for managing the risk of android ransomware," cited By 0, 2022, pp. 773–779. DOI: 10.1109/ICDABI56818.2022.10041528. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85149284732&doi=10.1109%5C%2fICDABI56818.2022.10041528&partnerID=40&md5=494c0f6a9a6ad84a51f7bcf8df806515.

[34] A. Atapour-Abarghouei, S. Bonner, and A. S. McGough, "Volenti non fit injuria: Ransomware and its victims," in *2019 IEEE International Conference on Big Data (Big Data)*, 2019, pp. 4701–4707. DOI: 10.1109/BigData47090.2019.9006298.

[35] U. Javed Butt, M. Abbod, A. Lors, H. Jahankhani, A. Jamal, and A. Kumar, "Ransomware threat and its impact on scada," cited By 12, 2019. DOI: 10.1109/ICGS3.2019.8688327. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85065013177&doi=10.1109%5C%2fICGS3.2019.8688327&partnerID=40&md5=1d51d3858d4d5f58c83c78b033cf756f.

[36] E. Galinkin, "Winning the ransomware lottery: A game-theoretic approach to preventing ransomware attacks," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 13061 LNCS, pp. 195–207, 2021, cited By 1. DOI: 10.1007/978-3-030-90370-1_11. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85119339316&doi=10.1007%5C%2f978-3-030-90370-1_11&partnerID=40&md5=c2cb05460294038978e342fc39b39ba9.

[37] H. Ghayoomi, K. Laskey, E. Miller-Hooks, C. Hooks, and M. Tariverdi, "Assessing resilience of hospitals to cyberattack," *Digital Health*, vol. 7, 2021, cited By 7. DOI: 10.1177/20552076211059366. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85120318755&doi=10.1177%5C%2f20552076211059366&partnerID=40&md5=9251c6c84017d7c23f67e3bcf498de92.

[38] G. Kim, S. Kim, S. Kang, and J. Kim, *A method for decrypting data infected with hive ransomware*, 2022. DOI: 10.48550/ARXIV.2202.08477. [Online]. Available: https://arxiv.org/abs/2202.08477.

[39] A. M. Maigida, S. M. Abdulhamid, M. Olalere, J. K. Alhassan, H. Chiroma, and E. G. Dada, "Systematic literature review and metadata analysis of ransomware attacks and detection mechanisms," *Journal of Reliable Intelligent Environments*, vol. 5, no. 2, pp. 67–89, May 2019. DOI: 10.1007/s40860-019-00080-3. [Online]. Available: https://doi.org/10.1007/s40860-019-00080-3.

[40] R. Fang, M. Xu, and P. Zhao, *Should the ransomware be paid?* 2020. DOI: 10.48550/ARXIV.2010.06700. [Online]. Available: https://arxiv.org/abs/2010.06700.

[41] N. Thamer and R. Alubady, "A survey of ransomware attacks for healthcare systems: Risks, challenges, solutions and opportunity of research," in *2021 1st Babylon International Conference on Information Technology and Science (BICITS)*, IEEE, Apr. 2021. DOI: 10.1109/bicits51482.2021.9509877. [Online]. Available: https://doi.org/10.1109/bicits51482.2021.9509877.

[42] D. Min, Y. Ko, R. Walker, J. Lee, and Y. Kim, "A content-based ransomware detection and backup solid-state drive for ransomware defense," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 41, no. 7, pp. 2038–2051, 2022, cited By 3. DOI: 10.1109/TCAD.2021.3099084. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85111605845&doi=10.1109%5C%2fTCAD.2021.3099084&partnerID=40&md5=fc4da7d5849de3ff9f4d74e55904f129.

[43] B. Reidys, P. Liu, and J. Huang, *Rssd: Defend against ransomware with hardware-isolated network-storage codesign and post-attack analysis*, 2022. DOI: 10.48550/ARXIV.2206.05821. [Online]. Available: https://arxiv.org/abs/2206.05821.

[44] T. Reshmi, "Information security breaches due to ransomware attacks - a systematic literature review," *International Journal of Information Management Data Insights*, vol. 1, no. 2, p. 100 013, Nov. 2021. DOI: 10.1016/j.jjimei.2021.100013. [Online]. Available: https://doi.org/10.1016/j.jjimei.2021.100013.

[45] A. Marget, *Snapshots vs. backups: Which is better?* Apr. 2022. [Online]. Available: https://www.unitrends.com/blog/snapshots-vs-backups.

[46] M. E. Ahmed, H. Kim, S. Camtepe, and S. Nepal, *Peeler: Profiling kernel-level events to detect ransomware*, 2021. DOI: 10.48550/ARXIV.2101.12434. [Online]. Available: https://arxiv.org/abs/2101.12434.

[47] R. Moussaileb, N. Cuppens, J.-L. Lanet, and H. L. Bouder, "A survey on windows-based ransomware taxonomy and detection mechanisms," *ACM Computing Surveys*, vol. 54, no. 6, pp. 1–36, Jul. 2021. DOI: 10.1145/3453153. [Online]. Available: https://doi.org/10.1145/3453153.

[48] C. Onwubiko, "Focusing on the recovery aspects of cyber resilience," in *2020 International Conference on Cyber Situational Awareness, Data Analytics and Assessment (CyberSA)*, IEEE, Jun. 2020. DOI: 10.1109/cybersa49311.2020.9139685. [Online]. Available: https://doi.org/10.1109/cybersa49311.2020.9139685.

[49] M. Shon, J. Park, H. Kim, K. Won, K. Park, and J. Hong, "A robust and secure backup system for protecting malware," cited By 0, vol. Part F147772, 2019, pp. 1432–1437. DOI: 10.1145/3297280.3297424. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85065669339&doi=10.1145%5C%2f3297280.3297424&partnerID=40&md5=01e466ea9e42e54d09d12b0897faccda.

[50] J. A. H. Silva, L. I. B. Lopez, A. L. V. Caraguay, and M. Hernandez-Alvarez, "A survey on situational awareness of ransomware attacks—detection and prevention parameters," *Remote Sensing*, vol. 11, no. 10, p. 1168, May 2019. DOI: 10.3390/rs11101168. [Online]. Available: https://doi.org/10.3390/rs11101168.

[51] V. Szucs, G. Aranyi, and A. David, "Introduction of the ards nti-ransomware defense system model based on the systematic review of worldwide ransomware attacks," *Applied Sciences*, vol. 11, no. 13, p. 6070, Jun. 2021. DOI: 10.3390/app11136070. [Online]. Available: https://doi.org/10.3390/app11136070.

[52] A. Tandon and A. Nayyar, "A comprehensive survey on ransomware attack: A growing havoc cyberthreat," in *Data Management, Analytics and Innovation*, Springer Singapore, Sep. 2018, pp. 403–420. DOI: 10.1007/978-981-13-1274-8_31. [Online]. Available: https://doi.org/10.1007/978-981-13-1274-8_31.

[53] S. Vasoya, K. Bhavsar, and N. Patel, *A systematic literature review on ransomware attacks*, 2022. DOI: 10.48550/ARXIV.2212.04063. [Online]. Available: https://arxiv.org/abs/2212.04063.

[54] C. Constantinescu and S. Seshadri, "Sentinel - ransomware detection in file storage," cited By 1, 2021. DOI: 10.1145/3456727.3463834. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85108453020&doi=10.1145%5C%2f3456727.3463834&partnerID=40&md5=1154c833c77c843554dc48e5923ee868.

[55] W. Lao, Z. Chen, B. Gao, J. Wang, Y. Tao, and R. Zhang, "Rap: Ransomware protection scheme based on blockchain," cited By 0, 2022, pp. 13–20. DOI: 10.1109/ICCECE54139.2022.9712682. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85126986213&doi=10.1109%5C%2fICCECE54139.2022.9712682&partnerID=40&md5=44d4a29de35d5248b72965e9903c0207.

[56] C. Lee and K. Lee, "Impact analysis of resilience against malicious code attacks via emails," *Computers, Materials and Continua*, vol. 72, no. 3, pp. 4803–4816, 2022, cited By 0. DOI: 10.32604/cmc.2022.025310. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85128618558&doi=10.32604%5C%2fcmc.2022.025310&partnerID=40&md5=e40e5909e51a10526f757b190ba0cbe6.

[57] Z. Manjezi and R. A. Botha, "Preventing and mitigating ransomware," in *Communications in Computer and Information Science*, Springer International Publishing, 2019, pp. 149–162. DOI: 10.1007/978-3-030-11407-7_11. [Online]. Available: https://doi.org/10.1007/978-3-030-11407-7_11.

[58] A. Kapoor, A. Gupta, R. Gupta, S. Tanwar, G. Sharma, and I. E. Davidson, "Ransomware detection, avoidance, and mitigation scheme: A review and future directions," *Sustainability*, vol. 14, no. 1, p. 8, Dec. 2021. DOI: 10.3390/su14010008. [Online]. Available: https://doi.org/10.3390/su14010008.

[59] S. Maniath, P. Poornachandran, and V. G. Sujadevi, "Survey on prevention, mitigation and containment of ransomware attacks," in *Communications in Computer and Information Science*, Springer Singapore, 2019, pp. 39–52. DOI: 10.1007/978-981-13-5826-5_3. [Online]. Available: https://doi.org/10.1007/978-981-13-5826-5_3.

[60] A. Melaragno and W. H. Casey, "Detecting ransomware execution in a timely manner," Jan. 2022. DOI: https://doi.org/10.48550/arxiv.2201.04424.

[61] A. Pagan and K. Elleithy, “A multi-layered defense approach to safeguard against ransomware,” cited By 2, 2021, pp. 942–947. DOI: 10.1109/CCWC51732.2021.9375988. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85103442486&doi=10.1109%5C%2fCCWC51732.2021.9375988&partnerID=40&md5=bf4e9e69dabea5c25ebff55a8c5070c4.

[62] A. Z. Abualkishik, A. A., and Y. Gulzar, “Disaster recovery in cloud computing systems: An overview,” *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 9, 2020. DOI: 10.14569/ijacsa.2020.0110984. [Online]. Available: https://doi.org/10.14569%2Fijacsa.2020.0110984.

[63] W. Song *et al.*, *Crypto-ransomware detection through quantitative api-based behavioral profiling*, 2023. DOI: 10.48550/ARXIV.2306.02270. [Online]. Available: https://arxiv.org/abs/2306.02270.

[64] A. Golev, R. Hristev, M. Veselinova, and K. Kolev, “Crypto-ransomware attacks on linux servers: A data recovery method,” vol. 21, pp. 19–29, Nov. 2022.

[65] S. Mujeye, “Ransomware: To pay or not to pay? the results of what it professionals recommend,” cited By 0, 2022, pp. 76–81. DOI: 10.1145/3520084.3520096. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85129143819&doi=10.1145%5C%2f3520084.3520096&partnerID=40&md5=2e7995431bffdf02c5658bd20c8e427e.

[66] M. Singhal, “Protecting customer databases to shield business data against ransomware attacks and effective disaster recovery in a hybrid production environment,” cited By 0, 2022. DOI: 10.1145/3590837.3590927. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85162263561&doi=10.1145%5C%2f3590837.3590927&partnerID=40&md5=8a26d82ee7b5111d40a92012329af1ac.

[67] Z. Li and H. Suo, “Research on cloud service reliability evaluation model from the perspective of equal protection 2.0,” *Journal of Physics: Conference Series*, vol. 1673, no. 1, p. 012 058, Nov. 2020. DOI: 10.1088/1742-6596/1673/1/012058. [Online]. Available: https://doi.org/10.1088/1742-6596/1673/1/012058.

[68] M. May and E. Laron, “Combating ransomware using content analysis and complex file events,” cited By 4, 2019. DOI: 10.1109/NTMS.2019.8763851. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85070419719&doi=10.1109%5C%2fNTMS.2019.8763851&partnerID=40&md5=0013f56d69ffb45ab505fd09a6709164.

[69] J. Castiglione and D. Pavlovic, “Dynamic distributed secure storage against ransomware,” *IEEE Transactions on Computational Social Systems*, vol. 7, no. 6, pp. 1469–1475, Dec. 2020. DOI: 10.1109/tcss.2019.2924650. [Online]. Available: https://doi.org/10.1109/tcss.2019.2924650.

[70] M. Baykara and B. Sekin, “A novel approach to ransomware: Designing a safe zone system,” in *2018 6th International Symposium on Digital Forensic and Security (ISDFS)*, IEEE, Mar. 2018. DOI: 10.1109/isdfs.2018.8355317. [Online]. Available: https://doi.org/10.1109/isdfs.2018.8355317.

[71] A. Lai *et al.*, “Ransomsoc: A more effective security operations center to detect and respond to ransomware attacks,” *Journal of Internet Services and Information Security*, vol. 12, no. 3, pp. 63–75, 2022, cited By 0. DOI: 10.22667/JISIS.2022.08.31.063. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85138160598&doi=10.22667%5C%2fJISIS.2022.08.31.063&partnerID=40&md5=1545e0b2a940de4b60337731fa215557.

[72] A. Hassanzadeh *et al.*, “A review of cybersecurity incidents in the water sector,” 2020. DOI: 10.48550/ARXIV.2001.11144. [Online]. Available: https://arxiv.org/abs/2001.11144.

[73] M. Tarka, M. Blankstein, and P. Schottel, “The crippling effects of a cyberattack at an academic level 1 trauma center: An orthopedic perspective,” *Injury*, vol. 54, no. 4, pp. 1095–1101, 2023, cited By 0. DOI: 10.1016/j.injury.2023.02.022. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85148351691&doi=10.1016%5C%2fj.injury.2023.02.022&partnerID=40&md5=47319a3058084e8ac548eb6f2bc4c178.

[74] C. F. NZ, *Don’t use disk mirroring as backup - data recovery*, Dec. 2014. [Online]. Available: https://www.datarecovery.co.nz/blog-posts/dont-use-disk-mirroring-as-backup/.

[75] ARCSERVE, *Can mirroring replace backups in your disaster recovery strategy? - arcserve*, 2014. [Online]. Available: https://www.arcserve.com/blog/can-mirroring-replace-backups-your-disaster-recovery-strategy.

[76] C. Scherb, L. B. Heitz, F. Grimberg, H. Grieder, and M. Maurer, *A serious game for simulating cyberattacks to teach cybersecurity*, 2023. DOI: 10.48550/ARXIV.2305.03062. [Online]. Available: https://arxiv.org/abs/2305.03062.

[77] J. Ahn *et al.*, *Key-ssd: Access-control drive to protect files from ransomware attacks*, 2019. DOI: 10.48550/ARXIV.1904.05012. [Online]. Available: https://arxiv.org/abs/1904.05012.

[78] J. Park, Y. Jung, J. Won, M. Kang, S. Lee, and J. Kim, “Ransomblocker: A low-overhead ransomware-proof ssd,” cited By 10, 2019. DOI: 10.1145/3316781.3317889. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85067820276&doi=10.1145%5C%2f3316781.3317889&partnerID=40&md5=159de107f0776ee1d9622f0dcb089e12.

[79] K. Begovic, A. Al-Ali, and Q. Malluhi, “Cryptographic ransomware encryption detection: Survey,” 2023. DOI: 10.48550/ARXIV.2306.12008. [Online]. Available: https://arxiv.org/abs/2306.12008.

[80] J. Kaur, *A secure and smart framework for preventing ransomware attack*, 2020. DOI: 10.48550/ARXIV.2001.07179. [Online]. Available: https://arxiv.org/abs/2001.07179.

[81] K. Lee, S.-Y. Lee, and K. Yim, “Machine learning based file entropy analysis for ransomware detection in backup systems,” *IEEE Access*, vol. 7, pp. 110 205–110 215, 2019. DOI: 10.1109/access.2019.2931136. [Online]. Available: https://doi.org/10.1109/access.2019.2931136.

[82] Y. Lemmou, J.-L. Lanet, and E. Souidi, “A behavioural in-depth analysis of ransomware infection,” *IET Information Security*, vol. 15, no. 1, pp. 38–58, 2021, cited By 9. DOI: 10.1049/ise2.12004. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85102649294&doi=10.1049%5C%2fise2.12004&partnerID=40&md5=4a1c09b68f036fbde81c7e16fdcfe3de.

[83] J. Morris, D. Lin, and M. Smith, *Fight virus like a virus: A new defense method against file-encrypting ransomware*, 2021. DOI: 10.48550/ARXIV.2103.11014. [Online]. Available: https://arxiv.org/abs/2103.11014.

[84] W. Wei, M. Qiao, E. Butler, and D. Jadav, “Graph representation learning based vulnerable target identification in ransomware attacks,” cited By 0, 2022, pp. 2423–2430. DOI: 10.1109/BigData55660.2022.10021008. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85147916940&doi=10.1109%5C%2fBigData55660.2022.10021008&partnerID=40&md5=c1877e09c55ec9ff16eb2a344152ff93.

[85] R. Fang, M. Xu, and P. Zhao, *Should the ransomware be paid?* 2020. DOI: 10.48550/ARXIV.2010.06700. [Online]. Available: https://arxiv.org/abs/2010.06700.

[86] A. Singh, M. Ali, B. Balamurugan, and V. Sharma, “Blockchain: Tool for controlling ransomware through pre-encryption and post-encryption behavior,” cited By 3, 2022, pp. 584–589. DOI: 10.1109/CCiCT56684.2022.00107. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85141720326&doi=10.1109%5C%2fCCiCT56684.2022.00107&partnerID=40&md5=0281363f1f9e0d3a0d6b6046e6295f56.

[87] A. Ren, C. Liang, I. Hyug, S. Brohi, and N. Jhanjhi, “A three-level ransomware detection and prevention mechanism,” *EAI Endorsed Transactions on Energy Web*, vol. 7, no. 26, 2020, cited By 8. DOI: 10.4108/eai.13-7-2018.162691. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85082646000&doi=10.4108%5C%2feai.13-7-2018.162691&partnerID=40&md5=8eca5a1161fb5075edc337847c80585a.

[88] Z. Li and Q. Liao, “Preventive portfolio against data-selling ransomware—a game theory of encryption and deception,” *Computers and Security*, vol. 116, 2022, cited By 3. DOI: 10.1016/j.cose.2022.102644. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85125853570&doi=10.1016%5C%2fj.cose.2022.102644&partnerID=40&md5=a1bb610d4ea7474e5adb98956cf0d1bf.

[89] K. Lee, J. Lee, S.-Y. Lee, and K. Yim, “Effective ransomware detection using entropy estimation of files for cloud services,” *Sensors*, vol. 23, no. 6, 2023, cited By 0. DOI: 10.3390/s23063023. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85151225279&doi=10.3390%5C%2fs23063023&partnerID=40&md5=52bc291bd9d09c3c63770b54dcb9c11f.

[90] G. F. O. of Information Security, *It-grundschutz*, 2022. [Online]. Available: https://www.bsi.bund.de/EN/Themen/Unternehmen-und-Organisationen/Standards-und-Zertifizierung/IT-Grundschutz/it-grundschutz_node.html.

[91] ISO, *Iso/iec 27001 standard – information security management systems*, Oct. 2022. [Online]. Available: https://www.iso.org/standard/27001.

[92] European Parliament and Council of European Union, *Regulation (eu) 2022/2554 of the european parliament and of the council of 14 december 2022 on digital operational resilience for the financial sector and amending regulations (ec) no 1060/2009, (eu) no 648/2012, (eu) no 600/2014, (eu) no 909/2014 and (eu) 2016/1011 (text with eea relevance)*, https://eur-lex.europa.eu/eli/reg/2022/2554/oj, 2022.

[93] T. Kumar and S. Kaur, "Cyber security in businesses: Challenges and recovery modes," in *2022 2nd International Conference on Emerging Smart Technologies and Applications (eSmarTA)*, IEEE, Oct. 2022. DOI: 10.1109/esmarta56775.2022.9935439. [Online]. Available: https://doi.org/10.1109/esmarta56775.2022.9935439.

[94] Cohesity, *5 ways ransomware renders backup useless*. [Online]. Available: https://www.cohesity.com/dm/tip-sheets/5-ways-ransomware-renders-backup-useless/.

[95] M. Miller, *Ransomware & the importance of offline backups » triaxiom security*, Triaxiom Security, Aug. 2019. [Online]. Available: https://www.triaxiomsecurity.com/ransomware-the-importance-of-offline-backups/.

[96] R. Chandramouli and D. Pinhas, "Security guidelines for storage infrastructure," Tech. Rep., Oct. 2020. DOI: 10.6028/nist.sp.800-209. [Online]. Available: https://doi.org/10.6028/nist.sp.800-209.

[97] LCH, *CYBER THREATS AND DATA RECOVERY CHALLENGES FOR FMIS*. Sep. 2021. [Online]. Available: https://www.lch.com/system/files/media_root/Cyber-Threats-and-Data-Recovery-Challenges-for-FMIs.pdf.

[98] Kaspersky, *Common TTPs of modern ransomware groups*. 2022. [Online]. Available: https://go.kaspersky.com/rs/802-IJN-240/images/Common-TTPs-of-the-modern-ransomware_low-res.pdf.

[99] M. ATT&CK, *Inhibit system recovery, technique t1490 - enterprise | mitre att&ck®*, attack.mitre.org. [Online]. Available: https://attack.mitre.org/techniques/T1490/.

[100] G. More, *Conti ransomware*, Qualys Security Blog, Nov. 2021. [Online]. Available: https://blog.qualys.com/vulnerabilities-threat-research/2021/11/18/conti-ransomware.

[101] J. Walter, *Lockbit 3.0 update | unpicking the ransomware's latest anti-analysis and evasion techniques*, SentinelOne, Jul. 2022. [Online]. Available: https://www.sentinelone.com/labs/lockbit-3-0-update-unpicking-the-ransomwares-latest-anti-analysis-and-evasion-techniques/.

[102] A. Wageh, *Lockbit ransomware analysis notes*, Medium, Aug. 2021. [Online]. Available: https://amgedwageh.medium.com/lockbit-ransomware-analysis-notes-93a542fc8511 (visited on 07/16/2023).

[103] N. Guillois, *Sodinokibi / revil malware analysis*, www.amossys.fr, Jul. 2020. [Online]. Available: `https://www.amossys.fr/fr/ressources/blog-technique/sodinokibi-malware-analysis/` (visited on 07/16/2023).

[104] Acronis, *Threat analysis: Babuk ransomware*, Acronis, Jul. 2021. [Online]. Available: `https://www.acronis.com/en-eu/blog/posts/babuk-ransomware/` (visited on 07/16/2023).

[105] KPN, *Tracking revil*, www.kpn.com. [Online]. Available: `https://www.kpn.com/security-blogs/Tracking-REvil.htm`.

[106] H. Newman, *Tape won't work for ransomware protection. here's why.* eSecurityPlanet, Sep. 2021. [Online]. Available: `https://www.esecurityplanet.com/threats/tape-wont-work-for-ransomware-protection/` (visited on 07/16/2023).

[107] A. C. S. C. (ACSC), *2023-03: Acsc ransomware profile – lockbit 3.0 | cyber.gov.au*, 2023. [Online]. Available: `https://www.cyber.gov.au/about-us/advisories/2023-03-acsc-ransomware-profile-lockbit-3.0`.

[108] J. Deyalsingh, N. Smith, E. Mattos, and T. Mclellan, *Alphv ransomware affiliate targets vulnerable backup installations to gain initial access*, Mandiant, Apr. 2023. [Online]. Available: `https://www.mandiant.com/resources/blog/alphv-ransomware-backup` (visited on 07/16/2023).

[109] K. Wojcieszek, S. Green, and E. Biasiotto, *Avoslocker ransomware update*, Kroll, Dec. 2022. [Online]. Available: `https://www.kroll.com/en/insights/publications/cyber/avoslocker-ransomware-update`.

[110] S. Sjouwerman, *Ransomware can destroy backups in four ways*, blog.knowbe4.com. [Online]. Available: `https://blog.knowbe4.com/ransomware-can-destroy-backups-in-four-ways` (visited on 07/16/2023).

[111] L. Tung, *Ransomware crooks hit synology nas devices with brute-force password attacks*, ZDNET, Jul. 2019. [Online]. Available: `https://www.zdnet.com/article/ransomware-crooks-hit-synology-nas-devices-with-brute-force-password-attacks/` (visited on 07/16/2023).

[112] cybleinc, *Cl0p ransomware: Active threat plaguing businesses worldwide*, Cyble, Apr. 2023. [Online]. Available: `https://blog.cyble.com/2023/04/03/cl0p-ransomware-active-threat-plaguing-businesses-worldwide/`.

[113] CISA, *Karakurt data extortion group | cisa*, www.cisa.gov, Jun. 2022. [Online]. Available: `https://www.cisa.gov/news-events/cybersecurity-advisories/aa22-152a`.

[114] S. Green and E. Biasiotto, *Black basta – technical analysis*, Kroll, Jan. 2023. [Online]. Available: `https://www.kroll.com/en/insights/publications/cyber/black-basta-technical-analysis`.

[115] M. Labs, *Mcafee atr analyzes sodinokibi aka revil ransomware-as-a-service - what the code tells us*, McAfee Blogs, Oct. 2019. [Online]. Available: `https://www.mcafee.com/blogs/other-blogs/mcafee-labs/mcafee-atr-analyzes-sodinokibi-aka-revil-ransomware-as-a-service-what-the-code-tells-us/`.

[116] Cybereason Global SOC & Cybereason Security Research, *Royal rumble: Analysis of royal ransomware*, www.cybereason.com, Dec. 2022. [Online]. Available: `https://www.cybereason.com/blog/royal-ransomware-analysis`.

[117] CISA, *Stopransomware: Cl0p ransomware gang exploits cve-2023-34362 moveit vulnerability cisa*, www.cisa.gov, Jun. 2023. [Online]. Available: `https://www.cisa.gov/news-events/cybersecurity-advisories/aa23-158a`.

[118] L. Abrams, *Massive qlocker ransomware attack uses 7zip to encrypt qnap devices*, BleepingComputer, Apr. 2021. [Online]. Available: `https://www.bleepingcomputer.com/news/security/massive-qlocker-ransomware-attack-uses-7zip-to-encrypt-qnap-devices/`.

[119] B. Toulas, *Qnap patches zero-day used in new deadbolt ransomware attacks*, BleepingComputer, Sep. 2022. [Online]. Available: `https://www.bleepingcomputer.com/news/security/qnap-patches-zero-day-used-in-new-deadbolt-ransomware-attacks/`.

[120] Cvedetails, *Rubrik : Security vulnerabilities*, www.cvedetails.com. [Online]. Available: `https://www.cvedetails.com/vulnerability-list/vendor_id-17599/Rubrik.html` (visited on 07/16/2023).

[121] Cvedetails, *Sophos : Security vulnerabilities*, www.cvedetails.com. [Online]. Available: `https://www.cvedetails.com/vulnerability-list/vendor_id-2047/Sophos.html` (visited on 07/16/2023).

[122] Cvedetails, *Unitrends : Security vulnerabilities*, www.cvedetails.com. [Online]. Available: `https://www.cvedetails.com/vulnerability-list/vendor_id-13281/Unitrends.html` (visited on 07/16/2023).

[123] Cvedetails, *Veeam : Security vulnerabilities*, Cvedetails.com. [Online]. Available: `https://www.cvedetails.com/vulnerability-list/vendor_id-15994/Veeam.html`.

[124] S. Gatlan, *Veeam fixes bug that lets hackers breach backup infrastructure*, BleepingComputer, Mar. 2023. [Online]. Available: `https://www.bleepingcomputer.com/news/security/veeam-fixes-bug-that-lets-hackers-breach-backup-infrastructure/` (visited on 07/16/2023).

[125] Datto, *Backup vulnerability: 4 targets hackers might utilize to infiltrate your backup solution*, Channel Futures, Mar. 2021. [Online]. Available: `https://www.channelfutures.com/from-the-industry/backup-vulnerability-4-targets-hackers-might-utilize-to-infiltrate-your-backup-solution`.

[126] C. Labs, *Cheerscrypt ransomware targets vmware esxi servers | cyware hacker news*, Cyware Labs, May 2022. [Online]. Available: `https://cyware.com/news/cheerscrypt-ransomware-targets-vmware-esxi-servers-d5f3f79a` (visited on 07/16/2023).

[127] Cvedetails, *Commvault : Security vulnerabilities*, www.cvedetails.com. [Online]. Available: `https://www.cvedetails.com/vulnerability-list/vendor_id-15702/Commvault.html` (visited on 07/16/2023).

[128] FOX-IT, *From backup to backdoor: Exploitation of cve-2022-36537 in r1soft server backup manager*, Fox-IT International blog, Feb. 2023. [Online]. Available: https://blog.fox-it.com/2023/02/22/from-backup-to-backdoor-exploitation-of-cve-2022-36537-in-r1soft-server-backup-manager/ (visited on 07/16/2023).

[129] V. Chebyshev, *It threat evolution q3 2019. statistics*, Securelist.com, Nov. 2019. [Online]. Available: https://securelist.com/it-threat-evolution-q3-2019-statistics/95269/.

[130] G. v. d. Klugt, *Cybercriminals exploit critical vulnerabilities in veeam backup*, Techzine Europe, Oct. 2022. [Online]. Available: https://www.techzine.eu/news/security/92401/cybercriminals-exploit-critical-vulnerabilities-in-veeam-backup/ (visited on 07/16/2023).

[131] A. Scroxton, *Ransomware gang exploiting unpatched veeam backup products | computer weekly*, ComputerWeekly.com, Apr. 2023. [Online]. Available: https://www.computerweekly.com/news/365535586/Ransomware-gang-exploiting-unpatched-Veeam-backup-products (visited on 07/16/2023).

[132] Ivanti, *76% of ransomware exploits pre-2020, report reveals*, Feb. 2023. [Online]. Available: https://www.ivanti.com/company/press-releases/2023/76-of-vulnerabilities-currently-exploited-by-ransomware-groups-were-discovered-before-2020-report-finds#:~:text=Old%5C%20is%5C%20still%5C%20gold%5C%20for.

[133] Cvedetails, *Qnap : Security vulnerabilities*, www.cvedetails.com. [Online]. Available: https://www.cvedetails.com/vulnerability-list/vendor_id-10080/Qnap.html (visited on 07/16/2023).

[134] Cvedetails, *Synology : Security vulnerabilities*, www.cvedetails.com. [Online]. Available: https://www.cvedetails.com/vulnerability-list/vendor_id-11138/Synology.html (visited on 07/16/2023).

[135] S. Ozarslan, *An underrated technique to delete volume shadow copies - deviceiocontrol*, 2021. [Online]. Available: https://www.picussecurity.com/resource/blog/technique-to-delete-volume-shadow-copies-deviceiocontrol.

[136] J. Gerend, *Wbadmin delete systemstatebackup*, Feb. 2023. [Online]. Available: https://learn.microsoft.com/en-us/windows-server/administration/windows-commands/wbadmin-delete-systemstatebackup.

[137] Microsoft, *Reagentc command-line options*, Aug. 2022. [Online]. Available: https://learn.microsoft.com/en-us/windows-hardware/manufacture/desktop/reagentc-command-line-options?view=windows-11.

[138] O. f. C. Rights (OCR), *Cyber security guidance material*, Jun. 2017. [Online]. Available: https://www.hhs.gov/hipaa/for-professionals/security/guidance/cybersecurity/index.html.

[139] S. Ranger, *Ransomware victims thought their backups were safe. they were wrong*, ZDNet, Feb. 2020. [Online]. Available: https://www.zdnet.com/article/ransomware-victims-thought-their-backups-were-safe-they-were-wrong/.

[140] Executive Order, *Executive order 13636 – improving critical infrastructure cybersecurity*, May 2013. [Online]. Available: https://obamawhitehouse.archives.gov/the-press-office/2013/02/12/executive-order-improving-critical-infrastructure-cybersecurity.

[141] N. Keller, *Getting started*, Feb. 2018. [Online]. Available: https://www.nist.gov/cyberframework/getting-started.

[142] Center for Internet Security, *Cis controls v8*, 2021. [Online]. Available: https://www.cisecurity.org/controls/v8.

[143] C. R. GmbH, *Digital operational resilience act (dora) - regulation (eu) 2022/2554*, 2022. [Online]. Available: https://www.digital-operational-resilience-act.com/.

[144] NERC, *North american electric reliability corporation critical infrastructure protection cyber security standards cip-002-1 through cip-009-1*, May 2006. [Online]. Available: https://www.nerc.com/pa/Stand/Cyber%5C%20Security%5C%20Permanent/Cyber_Security_Standards_Board_Approval_02May06.pdf.

[145] R. Awati, *What is nerc cip (critical infrastructure protection) and how does it work?* [Online]. Available: https://www.techtarget.com/searchsecurity/definition/North-American-Electric-Reliability-Corporation-Critical-Infrastructure-Protection-NERC-CIP#:~:text=The%5C%20North%5C%2f0American%5C%20Electric%5C%20Reliability.

[146] US Department of Defense, *Cybersecurity maturity model certification version 2.0*, Dec. 2021. [Online]. Available: https://dodcio.defense.gov/CMMC/Documentation/.

[147] A. C. S. C. (ACSC), *Information security manual*, 2023. [Online]. Available: https://www.cyber.gov.au/resources-business-and-government/essential-cyber-security/ism.

[148] Barracuda, *2023 ransomware insights*, Mar. 2023. [Online]. Available: https://www.barracuda.com/reports/ransomware-insights-report-2023.

[149] I. Cooke, *Backup and Recovery*. 2018. [Online]. Available: https://www.isaca.org/-/media/files/isacadp/project/isaca/articles/journal/2018/volume-1/backup-and-recovery_joa_eng_0118.

[150] H. Research, *BACKUP AND RECOVERY CHALLENGES AND TRENDS Prepared for Metallic, a Commvault Venture*. 2019. [Online]. Available: https://metallic.io/wp-content/uploads/2019/12/backup-and-recovery-challenges-and-trends.pdf.

[151] C. .-. D. S. Alliance, *Best practices in cybersecurity and cyber resilience A Data Security Alliance paper*. 2022. [Online]. Available: https://www.cohesity.com/resource-assets/white-paper/best-practices-in-cybersecurity-and-cyber-resilience-white-paper.pdf.

[152] Isaca, *Cobit | control objectives for information technologies | isaca*, 2007. [Online]. Available: https://www.bauer.uh.edu/parks/cobit_4.1.pdf.

[153] ENISA, *Boosting your organisation's cyber resilience - joint publication*, Feb. 2022. [Online]. Available: https://www.enisa.europa.eu/publications/boosting-your-organisations-cyber-resilience.

[154] N. S, *4 common backup problems and how to handle them*, MSP360, Jul. 2021. [Online]. Available: https://www.msp360.com/resources/blog/key-technical-backup-challenges-and-how-to-solve-them/.

[155] ISO, *Iso/iec 27001 standard – information security management systems*, Nov. 2005. [Online]. Available: https://www.iso.org/standard/27001.

[156] Clumio, *Overcoming the Challenges with Backup for AWS*. [Online]. Available: https://resources.enterprisetalk.com/ebook/CLUMIO-587-EN-1.pdf.

[157] VC3, *5 unexpected yet common data backup obstacles*, vc3.com. [Online]. Available: https://vc3.com/blog/5-unexpected-yet-common-data-backup-obstacles (visited on 07/16/2023).

[158] P. Solutions, *Discover the common backup challenges and solutions*, Prescient Solutions, Sep. 2019. [Online]. Available: https://www.prescientsolutions.com/blog/common-backup-challenges-and-solutions/ (visited on 07/16/2023).

[159] NAKIVO, *How to protect against ransomware attacks: 8 practices*, Nakivo, Aug. 2022. [Online]. Available: https://www.nakivo.com/blog/how-to-protect-against-ransomware-attacks/ (visited on 07/16/2023).

[160] N. S, *How to protect backups from ransomware*, MSP360, Sep. 2021. [Online]. Available: https://www.msp360.com/resources/blog/protect-backups-from-ransomware/.

[161] S. Cooper, *How to protect your backups from ransomware*, Comparitech, Jul. 2021. [Online]. Available: https://www.comparitech.com/net-admin/protect-backups-from-ransomware/.

[162] Veritas, *Immutable backups & ransomware attack mitigation*, www.veritas.com. [Online]. Available: https://www.veritas.com/information-center/immutable-backups.

[163] ISO, *Iso/iec 27002 standard – information security, cybersecurity and privacy protection — information security controls*, Oct. 2013. [Online]. Available: https://www.iso.org/standard/75652.html.

[164] L. Abrams, *Lockbit ransomware encryptors found targeting mac devices*, BleepingComputer, Apr. 2023. [Online]. Available: https://www.bleepingcomputer.com/news/security/lockbit-ransomware-encryptors-found-targeting-mac-devices/ (visited on 07/16/2023).

[165] K. Cole, *Move beyond the 3-2-1 rule for data backups*, TechBeacon, May 2023. [Online]. Available: https://techbeacon.com/security/move-beyond-3-2-1-rule-data-backups (visited on 07/16/2023).

[166] "Security and privacy controls for information systems and organizations," Tech. Rep., Sep. 2020. DOI: 10.6028/nist.sp.800-53r5. [Online]. Available: https://doi.org/10.6028/nist.sp.800-53r5.

[167] Cloudian, *Ransomware backup: How to get your data back*, Cloudian. [Online]. Available: https://cloudian.com/guides/ransomware-backup/ransomware-backup/.

[168] S. Mazor, *Ransomware recovery: The basics and 6 critical best practices*, bluexp.netapp.com, Jun. 2022. [Online]. Available: https://bluexp.netapp.com/blog/rps-blg-ransomware-recovery-the-basics-and-7-critical-best-practices (visited on 07/16/2023).

[169] W. C. Preston, *Ransomware: It's coming for your backup servers*, Network World, Dec. 2022. [Online]. Available: https://www.networkworld.com/article/3682659/ransomware-it-s-coming-for-your-backup-servers.html.

[170] M. University, *Response of Maastricht University to FOX-IT report*. 2020. [Online]. Available: https://www.maastrichtuniversity.nl/file/reponseofmaastrichtuniversitytofox-itreportpdf.

[171] ENISA, *Secure backups*, ENISA, Sep. 2021. [Online]. Available: https://www.enisa.europa.eu/securesme/cyber-tips/strengthen-technical-measures/secure-backups.

[172] Sophos, *Ransomware report: Sophos state of ransomware report 2023*, 2023. [Online]. Available: https://www.sophos.com/en-us/content/state-of-ransomware.

[173] D. Technologies, *Five steps to ransomware protection and recovery with APEX Backup Services A comprehensive guide to implementing an effective data resilience strategy*. 2022. [Online]. Available: https://www.delltechnologies.com/asset/zh-tw/solutions/apex/industry-market/steps-to-ransomware-protection-and-recovery-w-apex-backup-services-wp.pdf.

[174] W. C. Preston, *Tape backup as a defense vs. ransomware*, Network World, Sep. 2021. [Online]. Available: https://www.networkworld.com/article/3633934/tape-backup-as-a-defense-vs-ransomware.html (visited on 07/16/2023).

[175] H. Putta, *Top 5 reasons why backup and recovery in the cloud goes bad*, Microsoft, Nov. 2022. [Online]. Available: https://techcommunity.microsoft.com/t5/azure-governance-and-management/the-top-5-reasons-why-backup-and-recovery-in-the-cloud-goes/ba-p/3676072 (visited on 07/16/2023).

[176] Zenarmor, *The ultimate guide to data backup. best practices, methods, and storage options - zenarmor.com*, www.zenarmor.com. [Online]. Available: https://www.zenarmor.com/docs/network-security-tutorials/what-is-data-backup (visited on 07/16/2023).

[177] I. IT, *Threat spotlight: Lockbit black 3.0 ransomware*.

[178] J. Edwards, *Top 5 backup and recovery challenges and their remedies*, SearchDataBackup, Oct. 2020. [Online]. Available: https://www.techtarget.com/searchdatabackup/tip/Top-5-backup-and-recovery-challenges-and-their-remedies.

[179] D. Service, *Top 5 operational challenges in recovery management and how to solve them*, distribution-point.com. [Online]. Available: http://distribution-point.com/top5-operational-challenges-in-recovery-management-and-how-to-solve-them/ (visited on 07/16/2023).

[180] Commvault, *A Dialogue in Backup -Why Backup Breaks and How to Fix It*. 2018. [Online]. Available: https://cloud.kapostcontent.net/pub/db30db14-02a3-41fb-ae07-5a8907da341e/a-dialogue-in-backup-why-backup-breaks-and-how-to-fix-it?kui.

[181] A. Descalso, *What are the risks of not backing up data?* www.itsasap.com, Aug. 2021. [Online]. Available: https://www.itsasap.com/blog/why-backup-your-data.

[182] S. Hazlegreaves, *Backup and disaster recovery: Common encountered issues*, Open Access Government, Mar. 2021. [Online]. Available: https://www.openaccessgovernment.org/backup-and-disaster-recovery-common-encountered-issues/105449/ (visited on 07/16/2023).

[183] W. C. Preston, *Backup lessons from a cloud-storage disaster*, Network World, Apr. 2021. [Online]. Available: https://www.networkworld.com/article/3615678/backup-lessons-from-a-cloud-storage-disaster.html.

[184] A. Mallick, *Faq - protect backups from ransomware with azure backup - azure backup*, learn.microsoft.com. [Online]. Available: https://learn.microsoft.com/en-us/azure/backup/protect-backups-from-ransomware-faq (visited on 07/16/2023).

[185] N. C. S. Centre, *Mitigating malware and ransomware attacks*, www.ncsc.gov.uk, Feb. 2020. [Online]. Available: https://www.ncsc.gov.uk/guidance/mitigating-malware-and-ransomware-attacks.

[186] G. LaBrie, *Backup and recovery: Disaster plan challenges & tips*, blog.wei.com, Feb. 2017. [Online]. Available: https://blog.wei.com/backup-and-recovery-disaster-plan-challenges-tips.

[187] J. Naze, *Backup challenges*, Computerworld, Sep. 2004. [Online]. Available: https://www.computerworld.com/article/2812291/backup-challenges.html (visited on 07/16/2023).

[188] V. Consulting, *Backup challenges - ocala, gainesville, orlando*, Verteks Consulting, Inc., May 2022. [Online]. Available: https://www.verteks.com/2022/05/backup-challenges/ (visited on 07/16/2023).

[189] Arxys, *Backup and restore challenges in today's high threat environment*, Arxys, Jun. 2018. [Online]. Available: https://www.arxys.com/backup-and-restore-challenges-in-todays-high-threat-environment/ (visited on 07/16/2023).

[190] TechTarget, *A new/old weapon against ransomware: Tape backup*. [Online]. Available: https://www.techtarget.com/searchstorage/IronMountainCloud/A-New-Old-Weapon-Against-Ransomware-Tape-Backup (visited on 07/16/2023).

[191] Cloudian, *Storage tiering: Making the most of your storage investment*. [Online]. Available: https://cloudian.com/guides/data-backup/storage-tiering/.

[192] TechVera, *Backing up your data - which storage medium should you choose?* Jun. 2014. [Online]. Available: https://techvera.com/backing-up-your-data-which-storage-medium-should-you-choose/.

[193] A. Adshead, *What's the problem with nas backup? | computer weekly*, ComputerWeekly.com, Mar. 2020. [Online]. Available: https://www.computerweekly.com/feature/Whats-the-problem-with-NAS-backup (visited on 07/16/2023).

[194] S. Verma, *Servers & storage*, Servers & Storage, Mar. 2018. [Online]. Available: https://www.ibm.com/blogs/systems/best-data-backup-and-recovery-challenges-for-hybrid-cloud/.

[195] B. Jaylin, *Common issues with legacy backup and disaster recovery strategies*, OTAVA, Dec. 2021. [Online]. Available: https://www.otava.com/blog/common-issues-with-legacy-backup-and-disaster-recovery-strategies/.

[196] Deloitte, *Digital Resilience and Enterprise Recovery: Would your business survive a catastrophic cyber attack?* 2023. [Online]. Available: https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/risk/deloitte-uk-digital-resilience-and-enterprise-recovery-whitepaper.pdf.

[197] Loughtec, *Downtime: The real cost of ransomware*, LoughTec, Jul. 2022. [Online]. Available: https://loughtec.com/downtime-the-real-cost-of-ransomware/.

[198] I. D. M. (IDM), *58 percent of data backups are failing: Report | idm magazine*, idm.net.au. [Online]. Available: https://idm.net.au/article/0013354-58-percent-data-backups-are-failing-report (visited on 07/16/2023).

[199] E. Simmons, *News alert: Asigra highlights 5 data backup and recovery challenges associated with saas data*, The Last Watchdog, Jul. 2023. [Online]. Available: https://www.lastwatchdog.com/news-alert-asigra-highlights-5-data-backup-and-recovery-challenges-associated-with-saas-data/ (visited on 07/16/2023).

[200] Seagate, *Overcome 9 common challenges with cloud backup storage as a service*, 2022. [Online]. Available: https://www.seagate.com/resources/enterprise/ebook/seagate-overcome-9-common-challenges-with-cloud-backup-storage-as-a-service-ebook-final.pdf.

[201] S. Zhang, *6 biggest challenges in data recovery today*, Data Recovery Blog, Jul. 2017. [Online]. Available: https://www.datanumen.com/blogs/6-biggest-challenges-data-recovery-today/.

[202] A. Hurst, *Backup and recovery issues reported by 93 per cent of businesses*, Information Age, Apr. 2023. [Online]. Available: https://www.information-age.com/backup-and-recovery-issues-reported-by-93-per-cent-of-businesses-123503262/ (visited on 07/16/2023).

[203] Fortinet, *The 2023 Global Ransomware Report REPORT*. 2023. [Online]. Available: https://www.fortinet.com/content/dam/fortinet/assets/reports/report-2023-ransomware-global-research.pdf.

7
Appendix A: Informed Consent Form Template

This page is intentionally left blank to import the pdf file. Please follow to the next page for the informed consent form.

Delft University of Technology
HUMAN RESEARCH ETHICS
INFORMED CONSENT FORM

Participant Information/Opening Statement

You are being invited to participate in a research study titled Operational Resilience: Backup Strategies for Crisis Management in the Age of Ransomware. This study is being done by Dorukhan Yesilli, a Master's Student from TU Delft, in order to fulfill the requirements of their Master's Thesis. The participants for the interviews are selected from experts in their domain. This interview will take you approximately 30 minutes to complete.

Please note that your participation in this study **is entirely voluntary, and you can withdraw at any time.**

The purpose of these interviews is to recalibrate and support the conclusions that have come up during the literature review regarding ransomware strains and their attacks on backups, recovery strategies against ransomware through backups, and why such strategies fail or are limited in practical use. The potential outcomes are to generate a robust set of recommendations for organizations that face the threat of ransomware.

You will be asked a set of questions with regard to ransomware, backup and recovery strategies, and their limitations in real-life cases. You are free to not answer any questions without giving a reason.

The data obtained from the interviews will be used for creating technical summaries that are scoped from the interview recordings and transcripts to ensure that no personal data or personally identifiable information is present. The technical summaries will then be utilized in an aggregated way for the master's thesis to make conclusions about ransomware and backups.

To the best of our ability, your answers in this study will remain confidential. We will minimize any risks by ensuring that access management is correctly done on the data storage TU Delft Institutional Storage, where all data will be stored.

The personal data that will be collected and processed (for administrative purposes) are:

- Names and email addresses
- Audio/Video Recordings, Transcripts
- Signed Consent Forms
- Domain of Activity, Generic Job Description

Please note that by signing this informed consent form, you agree to the processing of the above personal data and that they will be archived and accessible with respect to the access management described in the further paragraphs.

There will be three pieces of data generated: audio recordings, transcripts, and technical summaries. Transcripts will be text versions of the audio recordings and will be pseudonymized. The technical summaries **will not** include any personal data. The technical summaries **are the only data** that will be published. The technical summaries will be generated by Dorukhan Yesilli by scoping down the interview question answers while omitting any unnecessary or identifiable data.

The technical summaries that will be extracted from the interviews, which will have no personally identifiable information, will be sent to you within a week of the interview, and until the end of the interview study (07/07/2023), you have time to ask for a change or omittance of information in the technical summaries. Please note that **only the technical summaries will be published** (and thus made public) **in the thesis from 31/08/2023**.

Access to all data is only for Dorukhan Yesilli (corresponding researcher) and Yury Zhauniarovich (responsible researcher). **Only the technical summaries (that have no personally identifiable information or personal data)** will also be accessible by external advisors from Ernst&Young (EY), namely, Ruurd Boomsma and Lukas Willinge. After the master's thesis is published, the audio recordings will be deleted; however, the pseudonymized transcripts, signed informed consent forms, and technical summaries will be archived. The responsibility of all material will fall to Yury Zhauniarovich, and the transcripts will be archived in a private instance of TU Delft Institutional Storage which is only accessible by Yury for two years starting from the moment that the thesis is published. The technical summaries, however, will be in a public repository appended to the master's thesis itself.

To mitigate any personally identifiable information or the sharing of sensitive information, the pseudonymization of the transcripts will be done, and only after that, the technical summaries will be created. The technical summaries will not have any identifiable information or any sensitive information. Afterward, as described before, the summaries will be sent to you, and you can point out if there needs to be any change or omittance. For unauthorized access or data leaks, proper access management will be carried out.

There will be no compensation paid to the interviewee.

You could ask any questions that come up before and throughout the interview.

You can also contact us below after the interview for any questions or concerns.

Contact details for the Corresponding Researcher:
Name: Dorukhan Yesilli
Email: h.d.yesilli@student.tudelft.nl

Contact details for the Responsible Researcher:
Name: Yury Zhauniarovich
Email: y.zhauniarovich@tudelft.nl

By signing this consent form, you signify that you are now aware of and understand everything that is laid out in the text and that you approve and consent to being interviewed.

Please continue below for explicit consent points applicable to this research.

Explicit Consent points

PLEASE TICK THE APPROPRIATE BOXES	**Yes**	**No**
A: RESEARCH PUBLICATION, DISSEMINATION, AND APPLICATION		
1. I agree that my responses, views, or other input can be quoted anonymously in research outputs	□	□

Signatures

______________________________ _____________________________ _________

Name of the participant [printed] Signature Date

I, as a researcher, have accurately read out the information sheet to the potential participant and, to the best of my ability, ensured that the participant understands what they are freely consenting.

____________________________ _____________________ _________

Researcher name [printed] Signature Date

Study contact details for further information:

Please feel free to contact me if you have any concerns or questions.

Contact details for the Corresponding Researcher:
Name: Dorukhan Yesilli
Email: h.d.yesilli@student.tudelft.nl

Contact details for the Responsible Researcher:
Name: Yury Zhauniarovich
Email: y.zhauniarovich@tudelft.nl

8 Appendix B: Semi-Structured Interview Protocol

In your organization, you work with, or previous work experience, please answer the following:

1. Can you tell me about your experience regarding ransomware, backups and challenges in recovery?
2. Are you knowledegable about ransomware attacks directed at backups?

8.1. Backup & Recovery Policy Exploration Phase

1. Do the organization(s) you work with have backup strategies, and what do they entail?
   (a) What are some complexities regarding data backups and their recovery you see within organizations?
   (b) Do the organization(s) think and establish backup & recovery frameworks in-house or are they inherited from third-party vendors of which their services based on?
   (c) (Optional) What data backup solutions have you seen utilized previously by organizations? Why?
   (d) In what ways are the backups protected in organizations?
   (e) Do you have any officially published framework (or legislation, etc.) you follow on an organizational level for backup strategies?
   (f) (Optional) Does (do) the organization(s) establish criticality levels for different groups of data? Is it okay to lose a certain portion of the data?
   (g) (Optional) What type of backups are created? (e.g., Full Backup, Incremental Backup, or Differential Backup)?
   (h) (Optional) How often are the backups tested for being fit for recovery?
   (i) (Optional) What is the retention period for backups?
   (j) (Optional) How often are backups created?
2. What are the recovery processes and procedures in the organization(s) that you work with?
   (a) (Optional) Do you follow any officially published frameworks regarding recovery, such as NIST CSF, ISO, or NIST SPs?
   (b) (Optional) What are some performance metrics that you follow for recovery? Can you elaborate? What are the RPOs (Recovery Point Objectives) and RTOs (Recovery Time Objectives)?

8.2. Limitation of the Backups and Impossibilities in Real-Life Practice

1. (Optional - If they have experienced a ransomware attack) How long has recovery taken from a ransomware attack?

   (a) At what stages during recovery have you encountered a bottleneck?

2. In your experience, what are the main challenges to recover from backups?

   (a) (Optional) How do factors like the size of an organization, the complexity of the IT Infrastructure, or industry-specific requirements may impact recovery from backups?

   (b) (Optional) How big is the threat of backups not being usable in a crisis due to factors such as being outdated, broken/corrupted, or encrypted by ransomware, etc.? (e.g., inadequate backups or testing, backup targeting attacks, analyzing previous backup infiltration)

   (c) (Optional) Could you elaborate on specific technical or operational challenges for organizations recovering from backups?

   (d) (Optional) What are the critical lessons learned for recovering/ or failing to recover from backups during times of crisis?

3. (Optional – If they have multiple tiered storages such as air-gapped (both physical and virtual) off-site backups) What are the challenges, limitations, and trade-offs experienced due to safer but more complex backup strategies that you utilize?

9 Appendix C: Technical Summaries of Interviews

9.1. Interview 1

Organizations may have regular penetration testing on their systems to simulate specific ransomware attacks. This simulation is done in order to understand whether specific ransomware families can encrypt their data or if their systems can detect it. This testing also involves the network and whether there can be access from the main network to the backup network. Bigger organizations tend to have redundant backup systems, which go into play the moment the main system is down. The critical point is access management must be set up properly, and only admins must be able to access backup systems.

Ransomware attack patterns are specified by the group that the family of the malware belongs to. For specific attacks against ransomware, these groups tend to target offsite backup storage by infiltrating the systems that create the backups. Then, they try to put malware on the storage media or encrypt the storage media itself so that it does not work when it is overwritten. For the backup network, ransomware groups tend to try identifying hosts that are responsible for the connection between the normal network and the backup network. Then they try to increase privileges by accessing a high-level IT admin, backup manager, or related account. If the network is also managed by the active directory and the group has admin privileges, they basically have full control. To crack these passwords, ransomware groups may utilize keylogging in several ways.

For more traditional methods, such as tape systems, ransomware attacks tend to attack the system that creates those tapes. One method is to modify certain USB drivers. This way, the USB driver can write and shows the system operation as normal, but it doesn't, in reality, write anything of utility but garbage to the tapes. For example, if one has privilege over the system, they can modify the USB drive to their version, which gives garbage out, and when restoration is initiated during a crisis from the tapes, the backups will not be worth anything.

It is important to note that the snapshots of the running system that has malware on it and is not detected will be terrible news after the recovery. The ransomware group then will already have full access to the systems after recovery. Backing up does not remove infections. The ransom attack can repeat until the victim pays.

Regarding ransomware increasing the difficulties of recovery, it is always difficult to restore systems from 0, and it really is situational based on how backups are configured in organizations. Most organizations tend to have both backups of data and also backups of systems. The data backups sometimes also include information about the operating system and users, so not only information about the database but where the backups are restored. Ransomware groups who make these attacks can then adjust their targets in the system and prefer tactics such as the payload of the ransomware being executed on certain types of runtime, user login, or other operations. Network sanitization in such an environment would then be very hard.

Established frameworks in the cyber field are NIST and CIS Controls. These have rules for backups and strategies that organizations could utilize.

Another big challenge in organizations is testing. If organizations do not know if a backup image works or not, and it is indeed corrupted, then during recovery, they have nothing. Moreover, if there is a version update for security for any type of third-party library or system component etc., and there are old version dependencies, there may be problems. A problem could be that necessary data would not be in the right spot or that there would be a need to migrate a lot of information which costs resources and would not sit well with organizations. Another big part is that organizations may not utilize cloud environments and that employees may be utilizing local copies on their personal laptops. Then, if an infection occurs, there are no backed-up files in the cloud.

Organizations should establish criticality metrics for their systems. First is the active directory, which is the service responsible for access. This system constructs file access, system access, etc., and keeps all the passwords and privileges of the user account. If the active directory is compromised and not running, then users cannot log in or access files. Other criticality metrics are situational and are dependent on the IT architecture of organizations. Here, a lot depends on business continuity plans and what data is important.

Organizations may choose to utilize third parties to set up their backup and database infrastructure. In-house development usually costs a lot of time and resources. Depending on the sector, organizations may have different mandates to comply with. To ensure compliance, they must be audited. For audits, they must specify their attack targets so that they can realize a thought-through risk management plan. There are, however, no specific requirements known in the industry for ransomware and recovery. Here, no regulations nor mandates make sense as organizational behavior change is quite complex and is managed through a lot of middle management. This makes it so that even if they know what risks may arise in their systems, there is a possibility that the problems that have been deemed critical for cybersecurity may not be fixed for a prolonged period.

Many organizations tend to see IT as a cost and not as a win. As long as they are not hit by a ransomware attack, an organization will not benefit from enhanced cybersecurity practices. It is also known that currently, there are insurance companies giving out insurance against ransomware attacks. If organizations see that the cost of maintaining and securing or even making decisions about these issues takes focus and resources away from their core value offering and/or profit-generating products, they could note that the marginal cost of not paying the ransom is bigger than just paying the ransom. This business thinking is well known and has made headlines; for example, Facebook broke many data protection laws and made the general news. However, to them selling the information was more profitable than the fine they paid due to GDPR.

The challenges that operations face during recovery are entirely dependent on how hard the ransomware groups attack their systems. If they have access to the active directory and the systems are infected in the organization, they are in trouble. If all the network and the laptops are encrypted, there also has to be a lot of stakeholder management, IT equipment management, etc., which would take a lot of time for all the users of the system to be in operation again. Another bottleneck could be not having recovery process guidelines. It is known from experience that many managers lack knowledge when it comes to recovery and incident response. It is important for organizations to have a sort of awareness to be able to reply to ransomware attacks under pressure.

Lastly, critical lessons learned for organizations could be that accountability and proper access control configurations are essential to ensure cyber hygiene. For example, an account that should have lower privileges might have been introduced as permission to override a password. These things may be forgotten and overlooked. Many organizations also tend to put bandaids to provide temporary fixes. However, temporary solutions tend to be always made permanent. For example, automatic backups may be a quick and easy solution; however, if the script that creates those automatic backups is rewritten, then they might actually start destroying the backups. The main issue is that there are not enough resources or the allocation of these resources to fix such issues. One IT person cannot manage a company-wide network. There must be a change in the understanding that cybersecurity is a cost.

9.2. Interview 2

Organizations need help with regards to creating incident response runbooks in order to know when to act and how to act against a ransomware attack. Here, it is important for an organization to know its backup capabilities and things to be aware of. Many organizations are immature with regard to their cybersecurity practices. This immaturity means that there is no network segregation, no network separation, or segmentation. In many cases, backups are located where the primary data is. This makes it so that under a crisis situation where the primary data are inaccessible, the backups are as well. An ideal setting with regards to this could be to utilize solutions such as the Veeam and Microsoft Azure Blob storage. Data is copied through Veeam and stored at the Blob storage. Then it’s all segmented.

There is not one backup strategy that can be defined as the best of all, but it is all based on a combination of practices and organizational capabilities.

Cybersecurity frameworks that gather the state-of-the-art are the ISO 27001, ISO 27002, COBIT, and the NIST Cybersecurity Framework. Although there is overlap between these frameworks, they are noted as the industry’s best guidelines. Basically, for backup strategies, backups should be stored at different locations, and proper access management must be conducted so that only certain people are able to access the backups. There also must be a redundant backup storage location so that if one fails, the other can always jump in to replace it.

When it comes to ransomware that specifically targets backups, it is mostly about privilege escalation and lateral movement to take over the system, delete files, delete logs, etc. Here, it is important to be able to get a high-level, holistic overview and approach to recommending suggestions.

Regarding backup solution providers, a big important factor is integration. For example, if an organization utilizes Microsoft Office 365 and that everything is already cloud-based, the organization would be making use of Microsoft licenses and solutions. Then it would be logical to follow suit and utilize the Microsoft Azure Blob storage as well. Then Microsoft guidance on creating and maintaining backups will be essential. With this, organizations would need one vendor for accountability, and that vendor would be responsible for installation, implementation, and configuration. However, if organizations may utilize multiple vendors such as Dell, IBM, Microsoft, and Amazon, then they would have a very segmented IT landscape that might not integrate well or intertwine with each other. Then, for example, in a backup and recovery situation, some of the programs may not communicate with each other. Organizations should thus try to stick with vendors that they already like and utilize the solution. However, this is still respective to organizational plans and would differ per circumstance.

When the backup strategies are analyzed, access management and control is a key issue. Who can view, alter or transfer the backup data, etc., are quite important questions to always monitor and ensure that the minimal level of access is given out to specific users like admins and backup managers. That is also part of the network and access control segregation. Moreover, organizations could ensure network separation, store their backups in a different location, and ensure that they are offline and offsite. It is also critical to ensure that recovery should be doable within a day so that the data loss is kept to a minimum. It should be noted that backups are intertwined with many domains within cybersecurity. Organizations usually tend to establish criticality measures for their data. They understand what data they handle and typically have data classification methodologies defined. They already have classifications done on their crown jewels. The crown jewels are the most critical assets, the most critical pieces of data that organizations would like to have recovered as soon as possible when an outage occurs. It is important, when dealing with backup strategies that the critical data are always prioritized. These backups must be tested for recovery, sometimes data as big as 100 terabytes to maybe a few petabytes. During recovery, the system is built off slowly until the whole environment is up and working again. Testing plans could be for example, to see if an organization may utilize a 10 Gigabit optic fiber network to recover the backup back online within 4 to 24 hours. Organizations, here again, must be aware of their capabilities and their wants. If they have the capability of restoring ten terabytes of data every hour, but their critical assets are 1 petabyte, this will take 100 hours. Thus, 100 hours before critical data are fully recovered. Moreover, the testing of backups is not done often enough. Excluding the financial sector, which is bound to mandatory requirements, many public and private sector companies typically do not have good backup strategies. The public and private sector companies are aware that they have backup capabilities, and they have backups; however, they may still not be able to recover from a ransomware attack. The reasoning could be simply that there is too much data, which are sensitive IP data etc., from a very long time ago. Many organizations still do not have very formalized and well-defined backup strategies and policies. Testing is also done on a very ad hoc basis or is minimal. The more critical data is, there must be longer retention period and a higher number of testing. Organizations must make sure that their teams are able to respond to crises and act.

Regarding challenges in cyber recovery, in this current era of information security, organizations have too much data, and their capabilities cannot match up to it. The world is very data-driven, and technology is trying to catch up. Organizations at the moment are always thus 1 step behind when it comes to data backup and data recovery. There is also the case of data corruption or old data formatting. These are the most common issues that are encountered within the industry. For example, an organization has old formatting, and perhaps backups were created with the FAT32 format. Currently, most file systems use NTFS. Then these are incompatible. Corruption, on the other hand, may occur due to a small network outage or variations in signal strength, etc. Then the data is deemed unreadable and unexecutable. Another trend that is seen is that many organizations do not consider their cybersecurity practices in-house. The overhead is deemed too significant for the size of the organization. Therefore, they opt for their IT to third-party service providers. However, these providers may not have the capabilities to adhere to security requirements. These third parties that are relied on might also be very slow in response. Organizations then, with this outsourcing of IT, do not have control over their data anymore, and this makes it very dangerous. The organization cannot have insights into what the IT service provider does with their data. This is a great struggle. In summary, there is too much data and too much work that needs to be done. The overhead of IT is significantly too high for organizations to invest in it. There is a lack of professionals who are schooled and would like to work in IT for this type of work. This then creates an incentive for companies to instead be more willing to pay the ransom.

Regarding the willingness to pay, let's create an example. If there are three experts in the IT team in-house, from the baseline of 50,000 euros a year, it will mean 150,000 euros each year to pay for the IT team. However, outsourcing IT will cost 30 to 40 thousand a year for all the capabilities and resources. This is the financial overhead that organizations would like to cut. They can pool this money to their core value offerings. This is especially true for organizations that are not IT oriented or focused. Thus, it is these differences that change the maturity levels of an organization.

Regarding critical lessons learned in the industry, a lot of organizations tend to say one thing but act differently. For example, they note that they do backup and recovery testing, but if their actions are actually analyzed, the latest backup and recovery test was done four years ago. Many organizations, as mentioned previously, do not simply have the capabilities or the resources. They do not have the time to push for backup recovery. Sometimes, the organizations are too big to do a full-scale backup and recovery. There are then three golden pillars in cybersecurity: people, processing, and technology. For people there are not enough skilled, experienced people; on the side of processing, they may have defined processes but don't follow up on them. There is no one to check and monitor. As for technology, they may have capabilities, or they might not.

9.3. Interview 3

Ransomware shines in the last step in the cyber kill chain. Previously, ransomware attacks were automated but also yielded limited results in terms of people actually paying for the ransom because the backups were recoverable, etc. These days, attacks are very manual. Many ransomware groups don't even need to encrypt the data; that is only the last step. The idea is that they try to get through the whole network, which is through manual interaction most of the time. Then they try to extort organizations through any means necessary. At first, the ransom is made on the encryption (and thus continuation of operations), and secondly, the ransom is made on the exfiltration of data. Ransomware groups do not tend to bet on only one horse. If a company can recover, the ransomware groups are less likely to get paid, which is their only incentive. Thus, ransomware groups attack backups in mostly a manual manner. In order to do this, they look through the network to discover where the backups are. In most cases, backups are quite easy to find. They can make this discovery as, at that point, they usually have the highest privileges that they can obtain, and they practically go and destroy the backup files. There are some big backup software providers who use standard Windows backups; the ransomware tends to target them as well. Ransomware groups tend to use public scripts to break the passwords to get credentials of high-privilege accounts in the system, and they just go through the console and basically wipe everything. Even if they fail in privilege escalation or if their commands do not work due to antivirus intrusion, they open up PowerShell and then wipe the disk partitions or do things like this. The ransomware attacks go really deep manually to destroy as much as possible.

As mentioned, ransomware shows its characteristics in the end. However, it is just like any malware. It can be detected by most defense systems normally, but at the point of detection, there is a high probability that your systems are already compromised. The encryption, deletion of backups, and ransom note is the last blow they give you to make it harder to recover. Ransomware infiltrates the network in basic ways, such as exploiting an unpatched vulnerability, etc., and that is what is generally not understood. People tend to think that ransomware is a thing on its own, but most of these attacks are done manually by groups in the system, and they make decisions when they are within the network. They do lateral movement, and they go up the main controller. They try to get privileges by accessing accounts with admin permissions. Once they do this, they remain hidden, looking for the data. They will look for anything that is valuable. While they are in the system, they will do research on what organization they breached, how much money they make, and how much money they can ask. They do their research and then generally extract terabytes of data depending on the size of the company. They will do everything to look for data. They go to all the organization's network servers and extract everything they find. This usually takes a month but may take longer or shorter depending on the amount of data. Then they will do their encryption. At that point, they already have full admin privileges, and they can disable all the organization's defenses. They have everything they need. That is why ransomware is effective. In the past, there was a different approach to ransomware. The ones that could be decrypted had programming errors. The attack strategy was different; it was fully automatic and was to hit as many people as they could. Now the attack is more focused. They ask, "Whom do we have as a victim?" They breach vulnerabilities, and if there is a zero-day vulnerability, they will scan half the Internet and breach 30-40 companies instead of running an automatic attack and hoping it works. A smaller company is probably easier to breach than a bigger one, and that is the way it goes. The manual touch and the negotiation make the ransomware effective.

Ransomware tries to infect backups in the longer term. It needs to be kept in mind that ransomware attacks don't happen in a small time frame. The attack is usually undetected for months at a time. The ransomware groups try to extend that period of how long they are in the network. For organizations, it is recommendable to keep data backups but wipe systems clean and rebuild from zero. As you restore a system backup that may already be infected with ransomware, then you basically give the attacker access to your network again, which is not ideal. Cloud backups may allow organizations a feasible solution for backups. If the ransomware group cannot have the credentials to it, they will just delete stuff from the cloud as much as they can. However, the cloud vendors are intelligent and may have backups that ransomware groups may not be able to remove from their systems. The cloud must also be well configured. Organizations then will be giving access to their data to the outside provider and will need to hope that they do not get breached at the same time. Regarding backup strategies and storage media, and corresponding attacks, tapes are a good solution as long as the tape management system access is secured properly. Digital backups are most vulnerable in terms of wipes as it is easier. The tapes are harder to erase, but there have been cases where ransomware can get access to the management portal, and they just tell the system to wipe the data as well. Here, it depends entirely on the strategy of the ransomware group. Systems and solutions, no matter which one you use, if implemented correctly, can be very effective. One common mistake organizations make is that the backup system is part of their domain. The backups need to be kept outside of the domain, in a fully separate system, for extra security.

If there are fully separate credentials and other policies, even if ransomware breaks your primary domain and gets admin privileges, their privilege will not be valid on the backup system. The backup system may have your credentials for your domain, but the primary domain must never have the credentials of your backups. Another thing is to make sure that the backups are immutable and to ensure that the backup system is not reachable through the primary domain.

A challenge that organizations face is that the IT systems must be monitored by humans. Even if all the defense systems are in place and the organization has everything up and running, the IT employees may not have monitored the alerts that their systems have been generating. However, Backups do not necessarily make IT systems more complex. It is only that their implementation needs to be done right, and that is where the most complexity comes in. However, once it is set up, it should be quite fine. One major thing that organizations face is the recovery of data. Most of the time, the backups are fine, but the organizations do not have any idea where to start recovering from or how fast it will go. This, indeed, is a question that organizations a lot of times cannot answer. Even if the concept is simple, to copy the data back to the system, it takes organizations, for instance, five days to get it to the data system, running the database and everything around it from backups. Another interesting observation regarding challenges in recovery is that many organizations do not know what their actual critical systems are. The rule of thumb is that if the whole network is encrypted, an organization needs to bring back about 80% of the revenue back online. Many organizations have an idea about what brings in money; however, they do not understand what the underlying systems are. That is typically seen very quickly in a crisis situation as they then start thinking about what they need, and then they figure out that some systems are dependent on others, and then you go from a small system component to a very big thing that must be recovered. That is where things go wrong. If an organization needs to recover 100 servers for full recovery compared to only 5 in order to get their core operations running, then they need to only restore 5. Many organizations, however, lack this awareness. The core IT operations are probably well known, but they need an application, but nobody knows what the actual application needs to do. For example, they need a database server, they need whatever data, and the moment they start uncovering those is the moment organizations get to the realization that they're missing half of the infrastructure to get them running.

Moreover, problems with backups happen quite often. There are sometimes corrupted backups, but more often than not, something is not backed up, rather than it not working. Sometimes organizations overlook things. Perhaps the system does not correctly sync. Typical IT problems happen over time. One critical observation here, then, is to ensure that backups and their corresponding procedures must be tested. However, this generally never happens. Many organizations just do backups and hope it works.

Regarding cybersecurity frameworks, the basics of what you need to look out for are the same; however, their application really changes with the dynamics of the company. It also matters how much the operations can be down for. If a company is losing millions, it could choose to just rebuild its systems from scratch and continue its operations. Utilizing system backups is not advisable as these backups may have been infiltrated by ransomware. Rebuilding the system from zero should be the basic approach to take. However, there is no company that can fully restore its operations without having any data. So you try to restore as much as you can from scratch. For infrastructure, this works very well, for example, in setting up a new mail server, setting up a new domain controller, etc. This type of stuff could be reinstalled without needing too much additional data from your previous setup. But, when you start your application, you have in the databases, the recommendation there, as noted prior, is always only to copy data and not copy the system snapshots or setup. Try to avoid system backups and try to only utilize data backups. Only if there is no choice organizations may use system backups, but they must be monitored well. The challenge is that even if there is a full forensics investigation on the system backups and nothing is found, the ransomware might still be in there. Even if the machine is not compromised in the system backup, due to perhaps something that is unpatched, etc., it may make your backups still vulnerable. Here, prioritization of what to restore becomes key.

9.4. Interview 4

Regarding ransomware attacks, the main thing is access. The main questions ransomware groups ask when they are in the network are, “What type of access does the attacker have on the computer, the Linux server, or the Windows system? Do they have access to different file shares? What is on this file share, and are backups saved here?” Then the idea is to remove, corrupt, or encrypt these backups.

Organizations tend to work with different backup software providers. Here, one of the biggest names is Dell. They offer a solution that is sort of a vault that opens automatically at one time and also closes 20 or 30 minutes later, and then the organization can push all the information and backups into that vault. Afterward, it is closed off and not integrated into the network anymore. It is a dedicated safe room within the organization that nobody can access. The airgap in the vault only opens for the uploading of the backup. Other backup solutions are that organizations if they have the resources, can physically own a department of service that is totally segregated from the normal network. In case of an attack, then they can shut down the primary servers. This is, however, indeed an expensive solution, as all of the hardware must be maintained and must be updated regularly. This also makes it so that backups must be tested and not corrupted. The third solution could be to rely on cloud providers such as Microsoft Azure. These cloud providers can make a backup. The organizations pay only for the amount of resources they use from these cloud providers. It is then easy to go by the playbook of these third parties regarding backup generation and maintenance. Main players in this market are Microsoft and Amazon for cloud services.

Regarding cybersecurity frameworks, there are some standards for cyber recovery. The NIST framework and the NIST 800 SP are good resources. There are also ISO guidelines for this. These are the most important guidelines to comply with. Many organizations have their internal policies for backups and recovery. They generate their own backups as they tend to want to be in charge of their own data. Therefore they have their own policies. They can, however, indeed integrate policies of established third parties if it is feasible and beneficial. The risk setting is crucial. What systems are important? What needs to be operational? Those are the main questions to answer with backups. Such assessments usually are done after buying the license for the product.

Organizations do try to comply with their established backup and recovery policies. Testing in this regard is crucial. Some organizations tend to have dedicated environments in order to test the backups to see whether they can rely on their backups. However, there are also many cases in which organizations do not test their backups. They note that they backed it up, but when something happens cannot recover from it.

In the face of a cyber attack, organizations need to have their critical infrastructure in their backups. The main thing then is to shut everything down and check whether a lot of sensitive information is gathered by the attackers. For the backups, just restart it in a dedicated environment that is isolated from the current network in order to stay operational for the most important function operations. There are challenges; however, if the critical and correct dependencies in the systems that are to be rebuilt are not properly established, backups might be missing essential information. Then you cannot rely on them. In this type of case, again, the testing phase is essential.

For backups, during an attack, organizations will mostly do a general scan to determine whether they have been infected or not. They also check their network activity and log books. However, this is also challenging as if someone had access to the Active Directory of the organization in the past, maybe three months before it infiltrated the backups 16 backups ago. Then monitoring this attack and ensuring that the attackers don't have access to the Active Directory on the recovered system is difficult. It is also important to note the way of the attack; what if it was an insider threat? Is that person still part of the organization? Perhaps they have access to more accounts? These are also things to note during recovery. Additionally, were there perhaps more backdoors in the systems even from a year ago? Attackers tend to hook themselves well into the systems, so organizations cannot remove them. For bottlenecks during recovery, the majority of the organizations are not mature enough to conduct risk assessments based on which system is more critical than other systems, but they think they are. But when it comes to backup and recovery, organizations integrate some backup or cyber recovery solutions but tend to miss some processes. There is also a lot of shadow IT in their environment, which can increase the difficulties of recovery. These shadows are not known, but many operations depend on them. Many organizations do not have a good map of dependencies and the relations between the systems laid out.

9.5. Interview 5

An organization moved from on-premise infrastructure to the cloud. One reason was noted to be the costs involved. Businesses try to cut costs whenever possible. Cloud innovation brings about that benefit in cost reduction. This mentality also comes into play during thinking about backups. Organizations tend to think about what data needs to be retained, how it will be secured, but also how it will relate to higher costs.

In a new environment, the organization focused on not previously utilized backup strategies but more cloud-viable strategies. Regarding frameworks utilized, it is important to note that organizations and the technology that they use work differently. Frameworks then become personal preference and what best suits the business and its continuity plans. Due to freshness, the organization utilized the practices that the third-party service provider noted as best practices. The first consideration that organizations who migrate their backup solution to the cloud make are: to decide on what is data sensitive, how sensitive the data is, and how long should the retention period for the backups be. In my case all data was regarded as sensitive, but the retention policies changed depending on the applications of the data. The storage policy changed with respect to whether it was employee data or whether it is customer data.

In this specific case, the organization used to utilize full backups; however, the cloud providers came up with incremental backups, which the organization thought was a good idea to save costs and to save storage space. The infrastructure was made better as time went on. Blunt moves, like moving from full backups to incremental backups, however, were attributed to the youngness of the resources and the skills of the information. They needed only to migrate the data fast and to start utilizing the cloud environment. This, however, changes from organization to organization. The organization did automatic tests every week to see if their backups were fit for recovery. Due to the previously described “all data is sensitive” approach, there were no differences in application of these tests. However, it is important to note that these tests would need to be altered to suit the criticality of the data. This also majorly depends on how often the data is being incrementally added to the backup. These are two measures that define how often a backup should be tested. Cloud providers also provide many tools and services that can ensure the safety and utility of the backups in order to allow operational continuity. One of these solutions is cross-regional backups. One [redundant copy of a] backup is then replicated in another availability zone. The cloud provider had created different availability zones and regions where organizations can replicate data across different regions and not just different availability zones. The availability zones are in the same region, so the organization put measures in place where there are two backups in different regions when it comes to data that is both sensitive and needs a large retention period.

The cloud providers also provide solutions for recovery. For example, a cloud provider Microsoft has their own solution for this. It could be a cloud-native solution, which many organizations would prefer to utilize. This might even be a service provided by a cloud provider without any extra costs.

Regarding incident response management etc., businesses are quite more or less siloed, so they want to operate individually. Each team within a company may be for themselves, so alternatives here may be so vast and different. Many teams focus only on what they're working on. Traditionally, these incidents are to be solved in-house and dealt with as soon as it happens. With the moving to the cloud, it becomes a bit more micro-service oriented. Where each team within the organization is for themselves, it is more like this than about the organization keeping to themselves and wanting to handle the whole process.

When incident response and backup and recovery plans are created, the information on performance metrics such as recovery time and recovery point objectives are selected. Therefore decision-making always follows a set of policies through the documentation of these selections. What organizations do also is create a specific set of operational methods for the cloud infrastructure. For example, there may be only three types of ways to create a backup. Someone cannot go into the cloud and create other backups with their specifications. Of course, these operational methods are based on the organization's needs. That is one way to enforce specific workflows and reduce deviations.

Regarding the backup strategies, the organization followed the 3-2-1 backup strategy during on-premise operations and other original frameworks; however, when they moved on to the cloud, they went with the best practices on the cloud. This was provided by the third-party vendor. The rulebooks were then established by Microsoft documentation, AWS documentation, and GCP documentation.

The main challenges regarding backups and recovery would be unusable backups, inadequate backups, or inadequate testing. More specifically, if there is no testing, then the organization will not know whether full recovery is possible when you need those backups. Testing is essential. On the other hand, data is not quite that straightforward, and it is very easily corrupted. A few mistakes or changing formats can cause huge failures on backups, so it needs to be treated sensitively. To be honest, on the technical part of things, when data is being handled and being transferred, etc., things must be taken a bit seriously.

One reason for the corruption of the backups may happen due to moving data. That is the biggest risk about data; when you are moving it and doing a migration from on-premise to the cloud, you're basically lifting and shifting information to the cloud, and that could be very dangerous. If it is a major cloud provider, it will provide the most secure ways to do this with no data loss. The organization moved to the cloud and tested the backups in the cloud before deleting the backups from on-premise. It must never be a cut and paste, but more of a copy and paste and then delete. Such small differences are quite important.

Another important thing is that even though the third-party cloud provider could be relied on with regards to supporting the organization in a time of trouble, some organizations have the business strategy to build a technology capability. Here, they try to educate their staff, train them to become knowledgeable and build skills to actually handle any problems that arise instead of relying on the third part. This, however, is not easy. Businesses always try to move things forward. There is no time to just talk right at that moment and to educate everyone or train everyone, but then they kick off in six months. This is not possible. What usually happens is that the team does not want to set up the IT infrastructure of the cloud but does not want to outsource it either. So here, then, the team and the third part do it together. The organization tries to build its own capabilities while utilizing, for example, Amazon services. For migrational purposes, Amazon Managed Services helps organizations to help move to the cloud. They provide engineers and people who are knowledgeable about the cloud to help you migrate. Thus, the organization team utilizes the experience of the Amazon team to get knowledgeable to get things going. It becomes a joint effort. Other challenges when it comes to the data are that many organizations do not even have an incident response plan in place, or it is very primitive. We recover from backups, but then whose responsibility is it to backup, or have the backups been tested? Are we able to recover? Because here, it is important to note that operationality is on different levels. You can set up transactions but then miss certain parts of the database and recover them later on. But these are not established. It is quite messy. It is because sometimes, when you are in the organization, you are mixed up. You say we have the backups, and they are taken care of. If anything goes wrong, when you come from the outside, it becomes even more confusing. This is important as even if the backups are there, if the protocols are not established that well to be dealing with these problems on the spot during a crisis is difficult. The organization might be very comfortable since they have the backups and say we conform to strategies or things like that; however, they need to be thinking more about running out the scenario, what would really happen during recovery, and whether we are totally safe, etc.

9.6. Interview 6

Many organizations follow up-to-date cybersecurity frameworks such as the DORA, which consists of requirements around backup and recovery that are immutable and are separated from the normal network, for example. Those requirements also come along in different security standards like the ISO. They utilize all things which come around, mainly related to disaster recovery, to be able to recover.

There is a lot of variety regarding backup solutions. Organizations backup in the cloud, they have a failover location, they use tapes, they use hard drives, they use Azure environments, they use third-party software. It depends on the organizations and their plans.

Organizations tend to aim to be resilient under all circumstances, which also means that they are able to recover from their own circumstances. If they are teething to go for something unreachable, they don't need immutable backups because if an attacker is never able to enter your system, you don't need immutable backups. So organizations do not aim for really achieving something unbreakable, but more so something which is always able to stand up again.

Organizations doing ransomware recovery in-house or whether they contact third-party vendors very much depends on the technology they are in. If it is a cloud provider, if it is SaaS, it is in-house, if it is on-premise. It depends.

Most organizations tend to/should test the backup procedures, and then you will also highlight the dependencies, and almost all organizations have their backup strategies determined on criticality. So coming from a critical business service, using the following applications, the following data for getting the certified label, and based on that, RTO and RPOs are defined. For ransomware recovery this is different, in case of a black swan event organizations can and will only recover the minimal services to be able to sustain as an organization. This has to do with the high cost and time related to these operations.

Although I do not know what the main challenges are during recovery, in setting up the backup plan, it is about the complexity of the application landscape. So for business project X, for example, to be able to handle large loads, you need a lot of different applications. So for the organizations to be able to determine those dependencies as well as the sequentiality in which these systems have to be rebuilt, creates a very complex issue.

It is important to realize there are two different moments in time when we talk about ransomware. The moment people infiltrate your system, and the moment they start encrypting. So an organization would want their data back one millisecond before they start encrypting. However, if the organization recovers the data back from that moment in time, the ransomware software or malware is still in it. So you want the data from just before the encryption, but you want to do an analysis on that data to be able to remove the actual malware from it. So it is not necessarily that the organization needs to go back to the moment ransomware entered their systems, but more so go to the latest data the organization has and then patch the leak.

Currently, depending on the technology, organizations tend to utilize all different types of backups. So for databases, incremental is very logical. If you look at Virtual Machines, they use snapshots which are full backups. This very much depends on the type of application that an organization has or the technology they are on. All these backup types could also go in a ransomware-proof solution, so they are not really relevant in that sense.

There are two main bottlenecks during cyber recovery. One is the analysis of what happened, and the second is rebuilding the complete infrastructure. For rebuilding, the organizations must know the ordering of what to rebuild and when. They need to have all the right information, from addresses for the main controllers and AAD information with regards to accounts, so they need a lot of different information, and it could be so that the predefined order may not even always be correct. Even if the responsibility and procedures are established beforehand, there will be some duration, time to spin up new machines, etc. So it is more that the complexity, technology, and different services make it time-consuming to rebuild while you still need to do the analysis.

Regarding critical lessons for organizations, the most important is to test your backups. If you haven’t tested, then you don’t know if they work. When the backup frequency is considered, the DORA framework says that minimal yearly, and then after every significant or material change, new backups must be generated. The scope, however, needs to be changed with respect to the organization. This change in scope must account for at least one full backup from a complete system or process restore if the organization would like to be prepared for ransomware.

The main frameworks that are recognized in the cybersecurity field are the DORA, NIST, ISO, and CIS controls. Other than that, the DB good practices are a good framework that is addressed for financial institutions.

To answer the question about change in employees resulting in lost experience and whether that contributes to challenges in recovery, although the case is quite specific, hopefully, the organizations document such procedures and policies. I do not think it is a probable situation.

9.7. Interview 7

One core backup strategy for organizations is the 3-2-1 backup strategy. There are three different backups, two different types of storage media utilized, which are also redundant on two separate locations, and then one is stored offsite. Of course, it is difficult to scale for large organizations. Testing these backups is also crucial, as well as testing the recovery from end-to-end so that the entirety of the recovery strategy is fit for use.

For analysis and forensics insights, the organizations must be careful not to recover before analysis takes place, as then evidence will be lost. This evidence is essential during potential litigations that will be faced in the aftermath of a cyber attack. Normally ransomware targets one system, then they move from one to another. However, if recovery is made without taking a forensic copy, then experts will not be able to understand the extent of the infiltration done by ransomware.

There are some organizations that focus more on “on-premise” solutions. However, the trend is that people are mostly moving towards a hybrid setup, where there is a part in the cloud, and there are parts on site. A lot of individual file hosting is moved to the cloud. OneDrive or other types of cloud storage are also more resilient against ransomware. In a case where the organization network is breached, they can go to the third-party cloud provider and ask them to roll back to a good state of the system.

In very large organizations, in-house capabilities are relatively limited. Many may tend to go to third-party providers who will deal with backups and recovery. This is especially true for incident response. The reason is that it is a very specific task, and having people in-house who deal with incident response would require resources and thinking into something like a very active threat model. The organization must know its high-value targets and critical data. For an average company, it is mostly not monetarily worth it to have your own in-house incident response teams; it simply is too expensive. Organizations such as banks, on the other hand, are secure and ensure that they have incident response teams in-house. Having them in-house is more convenient as they will know the infrastructure of the organization in more depth than a third party. One interesting thing to note here is documentation. Some organizations have barely documented network infrastructure leading to confusion regarding how things connect with each other, which makes analysis, backup, and recovery more difficult.

Many ransomware are quite sophisticated and run anti-forensics measures. They like to delete forensic artifacts or at least try to delete forensic artifacts, which makes them harder to find and analyze. Ransomware also tends to pivot to other systems that the third party who has come to investigate is not aware of. Perhaps, through this, they can redeploy their payloads again. But also, a lot of times, there are basically two different types of ransomware groups. Some of them just buy ransomware as a service, and they just deploy It somewhere, and they don't really do any anti-forensics. They just want to get the money as fast as they can. Some ransomware, on the other hand, is a little more sophisticated and wants to linger in the network for longer to get more critical information that they can threaten the organization with. These anti-forensics measures are difficult to handle but still doable through knowledgeable professionals.

With a big data breach, organizations must also comply with specific regulations such as the GDPR and do impact assessments. Here, bureaucracy regarding legal procedures becomes an issue.

Organizations analyze ransomware attacks usually in the aftermath of detection but prior to recovery. Here, they try to find important forensic artifacts and then just do some rudimentary analysis to send the results back to the main server. This shows the organization what systems have certain indicative compromises. Then the systems which need further investigation are analyzed. From there on, actual forensic copies are obtained, and then the analyzers are able to get deeper into the utilization of the network. Ransomware tends to upload data through the browser but sometimes uses more sophisticated methods. An analysis is also done on the network logs. Ransomware usually shows itself through patterns in the log. However, a challenge here is that companies tend to log only for a short amount of time, and by the time they realize they are hit by a ransomware attack, all the important network artifacts have probably already been overwritten, as these attacks take months. If the organization's network allows it, it can show how much data it actually leaves. Most of the time, however, organizations do not have this information. Many organizations have a retention period of 30 days for log data; however, this constitutes a lot of data and must be stored somewhere. The size of the logs then tends to be gigabytes to even terabytes of data depending on the size of the company.

To circumvent large data, the organization may try to use incremental backups. However, I think they are a little bit more susceptible to ransomware as the attackers can change everything. When you think about backups, you have, of course, the backups of user systems and also backups of servers and other more critical applications. There should be different strategies utilized for each of the different critical services, of course. Here, perhaps a good idea is to work with snapshotting. Regarding recovery, the infrastructure could be recovered quite quickly, and afterward, it is basically the end of it. The actual instance itself is relatively fast to solve; however, the long-term legal issues, analysis, etc., take more time. The longer-term considerations take about two months to get done. In honesty, it is the first couple of weeks which is quite hectic. Organizations try to see if the threat actor is actually out of their network. Afterward, they can rebuild, depending on how good their backups are; this time period can be faster or slower. If I were to guess, it is, on average, about 2 to 3 months for recovery from beginning to the end, from the actual impact to the recovery.

Some local ITs at organizations, in some cases, can also just restore from backups automatically through their technological utilities. However, there may be some communication issues if, for example, the organization has different locations across the country. Then the complexity of the organization makes it so that different IT teams or management may not be able to communicate effectively. Considering operational issues are critical at this period, this may prove a significant challenge. Perhaps the local teams may have different backup procedures and policies than the national level, responsible for their infrastructure. These things are all dependent on circumstances and the organization.

In a network sense, there are organizations that have different tier layers of the network and different subnets for different applications. These networks have differing policies and may have different retention periods for backups. Critical data backups must also have redundant copies, whereas less consideration may fall to lower priority data.

9.8. Interview 8

Organizations ask red teams to try to attack their systems, simulating ransomware. Ransomware typically gains access to the domain and then escalates privileges to exfiltrate the data. Those are the usual objectives of red teams. Sometimes there is also a white box approach where organizations are reviewed on how their backup procedures are and how they are protected. An objective is also to figure out how to access the backup system to see if it could be compromised. All these actions are typical of Ransomware attacks. The first stage always ransomware attacks is to remain stealthy and access the system. The access will then lead to taking control. If ransomware gets admin privileges in the system and has access to it, then it will be able to make a big impact. The ransomware then checks if they have access to backups and then learns the policies behind it. If there are offline backups, the ransomware needs to be able to compromise them in a way so that the backups and the network are still accessible. That is when they actually compromise the backups. This is more for long-term operations.

With regards to compromise, creating garbage backups is an idea, or the ransomware can poison the machines and stay under the radar. Then they would still have control and access to the system when the organization recovers. Here the ransomware attacker can use many different methods to create a back channel that gets the back control. Therefore, blocking the recovery option for the organization is not necessarily what they do; they just infiltrate the backups and let organizations recover. They will then get access back soon.

Regarding backup strategies in organizations, it is entirely dependent on the organization and how they are set up. However, for example, offline backups, even if it is called offline backups, are still connected in a way. That is the difficulty. How do you make it offline and ensure that nobody unauthorized can touch it? Even if there is a physical tape system, if they are automated, if the ransomware controls the tape system, then ransomware agents can just load the tape that they want to compromise if the organization would like to make automated backup systems in an offline system, that is usually difficult. If people need to do it manually, then the organization needs to make sure to follow up as it is done correctly as needed. Regarding frameworks, many organizations base themselves on industry standards, such as NIST and ISO, and on specific business-related policies that are set up internally. There are a lot of circumstantial decisions to be made that are dependent on the organization's characteristics. Some organizations may utilize Microsoft Azure and then have long retention periods, etc. The example would change per organization.

One important thing is to make sure that the organizations have people to recover quickly. The organizations lack quite a bit in the testing phase. It is good to have a backup, but when the time comes that the organization really needs to use it, usually there are people running around and saying, “What do we do?”. The testing phase is really critical in terms of backup strategy and in terms of business resilience. If you do not do that with the smaller companies, then it is really difficult to know what will happen in case you need it.

Protection measures against attacks are dependent on the company. However, the current trend is that organizations are moving more and more toward network segmentation. They do this finally after being told for 20 years that it is good practice. There is a lot of zero-trust lingo being thrown around. This is just basic security. Make sure that people can only get access to what they must, basically. Now organizations are more focused on identity access management, which is important as well. Because proper access controls are the basis of strong cyber hygiene; however, even with all security practices in place, organizations must assume that somebody will be able to get a strong access level to their network. Therefore, “How can they make sure that they only get access to what they are going to have?” is a critical question. We also now see more and more maturity in terms of detection in organizations, and now the shift is also being made to focus on business-specific rules. For example, if you get access outside of business hours and you do not follow normal business procedures, it could be something that can indicate an attack. With machine learning, we also see a big change in behavioral analysis.

The main recovery challenge, as described before, is definitely testing it. Make sure it works but also make sure that only the relevant data is backed up. Organizations do not need to back up the entirety of the operating system. For example, you just need the data. In the end, the size of the backup will just be the size of the data, which will only contain the most important data that the organization has. The OS, on the other hand, can easily be deployed from scratch, and then organizations can restore their data on top of it.

For not only ransomware but cybersecurity incidents, one important thing is that organizations must be prepared for such incidents. Everybody in the organization must know what to do in case of a crisis. Organizations should not wait for incidents to happen but be proactive in testing their processes and their crisis management. Who is responsible for what, what must be done by whom, who must be called, what, and where is the critical data? In general, organizations must be ready. Doing exercises to improve resilience is important. The only thing matters is how organizations react and how fast to incidents. It is impossible not to be breached in the current day and age. We see a shift in organizational thinking from the old “cybersecurity is a cost” to an essential part of business continuity.

9.9. Interview 9

Ransomware is a type of attack that tries to encrypt your critical data. The backups, on the other hand, are just one of the mechanisms to restore the data. Restoring data from a backup is a responsive action, and generating a backup is a preventive measure for the ransomware attack. When ransomware attacks are considered, it is an attack that firstly needs to be looked at from the perspective of how it can be prevented and, secondly, how it can be detected early. Thirdly, recovery requires good backups. Organizations must be helped in order to build early detection mechanisms for ransomware attacks. One recommended backup software provider is Veeam.

Ransomware tries to encrypt the data in the target system; however, there is a possibility that it can encrypt the backup data. Organizations must then ensure that their backup data are protected from being encrypted. The way the attack goes could be that software is downloaded on the network system at any point. At that stage, the ransomware is not different from any other malicious software. It cannot be discerned whether the malicious code is ransomware or something different. Here, only behavior-based detection is going to help with categorization. If it is ransomware, it will try to access multiple files in the system, and it will encrypt multiple files in your system. It will create some processes where their signature match and are linked to ransomware processes. When it encrypts, it will encrypt the files into some known ransomware file extensions. So a specialized system can check to see if these behaviors are apparent and alert the user immediately at runtime. There is also a different use case which is more complex but reliable in detection. If in the system there is more than the normal number of files that are being accessed, being changed into different formats, we can label it as a potential ransomware attack.

In regards to backups, the first step is to have your real-time data to be backed up, which is modified. Here, it depends on the system. If it is a critical system, then you might have a real-time backup going on. This usually happens in a different location on a different server. The backup procedures do not happen on the same server, as the ransomware runs on one particular server it tries to encrypt. Therefore while doing analytics, you can safely say the backup copy is on a different server. The second scenario is that usually, the backups happen periodically, once a day, usually at midnight, if it is not a highly critical system that does not need to be backed up in real-time. Then if early detection is achieved, the backup that happened last night is still safe, and the backup of today is only going to be done at the end of the day. Then again, the backup is also going to happen on a different system. From the design perspective of backup for best practices, if you succeed in lead detection, you can safely say that your backup is still safe. However, it is important to note that at the first stage, ransomware will still refer first to encrypting the first production data. Backups are stored on backup systems or even sometimes traditional methods such as tapes which the production system can also be stored in. However, here the biggest factors are that you have a good backup in a remote location that is safe and protected.

The backup strategies are made according to the value of the data. The value of the data is found during risk assessments that organizations utilize. They rate the data based on the CIA ratings, confidentiality, integrity, and availability. If the data is confidential and the availability requirements are high, then you would ensure that the data is backed up online with immediate replication. When you have sort of critical data, organizations use advanced backup mechanisms, which can also be hot and hot backup. This means that data is backed up the moment it is created. Depending on the architecture of the backup, you can store the data at the time of creation in a remote location. Later, a redundant copy could be made to tapes. There are also more advanced solutions that block backup copies from being altered by ransomware. Questions organizations must ask, however, are whether, even if they have good tools, do they allocate enough resources to the tools and whether they have a robust backup design. This notes whether the organizations do regular backups, the testing of the backups, and full recovery testing. In my experience, many organizations don't but claim that they have backups. Many organizations do not do end-to-end recovery testing. Among my long experience in the field, I have only seen a few organizations that were running backups with confidence and that, after every six months, switched to the other side of the DR system and tried to do restoration from backups. This decreases the effectiveness of the backups and is a main challenge.

For frameworks, many organizations start out by utilizing NIST. However, to my knowledge, NIST does not specifically focus on the backup side but says more about detecting anomalous behavior like ransomware. ISO 27000, ISO 27001, and NIST are actually very popular. Out of 10 organizations, nine would have based their policies on NIST.

One interesting thing is that organizations will now also need ransom negotiators. They are utilized to negotiate on the organization's behalf with the attacker to get the data back. An organization might try to build as robust of a strategy as they can, but they must always know that there can be a situation that they will get compromised. Thus, preparation in every aspect is key. The ransomware negotiator is also one of the aspects that make up a complete ransomware response statute. This, however, is not linked to willingness to pay but to how important the data is for you. Regarding challenges with respect to backups, there are different teams involved in an organization for IT infrastructure management. You have networking, an infrastructure team managing the OS file level, an application team, and a backup team. So the backup teams are usually closer to the infrastructure team, which means the team who builds the OS and manages the file system. So they try to do the backup at the file level, which works about 70% of the time. But about 30% there is an application that stores the business data. This application backup is different from the file system backup. If the applications, most of the time, are database oriented, they can be critical for the sequence of data that is stored. If you are storing files, you cannot truly determine in which sequence the data was stored at the moment the disruption happened. Therefore, they might be recovered at a file level, but that file is not enough to run the application because there were dependencies and sequential order.

One backup strategy case was that we stopped the application at midnight and took a backup. This is called an old backup from when the application stopped running. Then the application continues as normal. The data from the previous night onwards then is still in the files but not in the backup. This strategy ensures that every midnight we have a full backup, a cold backup that can be restored. This, however, raises the risk of having 24 hours of data loss, which was acceptable in our case. It is also a solid strategy in terms of the application. The application worked with was very sensitive to the data sequence. The data sequence had to be always maintained. The backup team and the file-level team would not know the sequence of this application. Therefore there would be missing information that the teams are not aware of. For such infrastructure teams, only OS-level file-level backups are good. This lack of awareness and narrow-sightedness of teams is one of the critical challenges observed.

Some organizations say that they back up their whole virtual machines. However, that is not enough to guarantee that the application can be brought back.

Linked to this, another issue is the coordination among different teams. Therefore when there is a disaster, there must be very good coordination among different teams: the backup team, the infrastructure team, and the application team. There are many things going wrong on the disaster recovery side. The backup team can have six resources where one may not have worked with the type of backup that was requested of them. Then the right processes were not done when it was in need. Disaster recovery needs detailed plans. If you have not done dry runs of the plan, it will not be executed successfully when the real crisis strikes. Even if there is a perfect technical backup and the sequencing issues are solved, if the human resources are not trained well, they will not be able to perform optimally.

9.10. Interview 10

Organizations tend to utilize some new emergent technologies provided by credible vendors such as Dell with their backup solutions. Third-party vendors can get the backup environment online, like flipping a switch. The vault solution basically recreates the environment of the company in a safe space, which is called a data vault. That is one of the latest trends in the field. The solutions indeed come with a bit of an investment. However, the results of these solutions may be very beneficial for organizations.

Regarding frameworks, organizations are recommended to utilize the DORA and NIST to gain insights on what to do in the event of a data breach or what to do if backups are affected and assess whether that falls into the context of the GDPR. The NIST framework, especially, is quite adhered to by organizations. Even though there are other frameworks than the ones mentioned, most conversations are had over the NIST and DORA frameworks. Of course, this depends on the context, sector, and business plans of the organization that is in question.

Regarding ransomware trends, nowadays, organizations are dealing with “big game hunting.” These are really sophisticated and done by organized crime groups like the Conti group or other state actors. They really specialize in a manual approach. We see very often that, they get initial access to the network through phishing. They get on to some of the endpoints and navigate their way through. They do reconnaissance about the network environment and stay under the radar to not get detected, so they stay in the network for quite some time. What ransomware actors try to achieve is to first identify the crown jewels of an organization. So “Where can the most precious assets be found? Data-wise and system-wise, and how can I jeopardize the organization as much as possible to extort the maximum amount of ransom?” are the questions that a ransomware actor thinks during the attack. Most of the time, the double extortion method and sometimes the triple extortion method are utilized. Triple extortion is when the attacker also reaches out to the customer of the victim organization to put more pressure. This is still one of the main ways that ransomware operates. This is a manual process. After this, it shifts into a more automated process where they will start with beaconing. Here what they do is they deploy as much C2C communication across the environment to automate the ransomware as much as possible. So that is a more automatic attack. Here they also try to reach the backups. Ransomware sometimes stays for such a long time in the environment that the organizations must think about the retention policies that are in place and think whether even the backups are still usable or whether they must be first cleaned before initiating recovery.

The infection of the backups, even the threat of it, is definitely one of the challenges. If the backups are affected, that basically means that the organization will most likely recreate its system from scratch. This costs resources and time. It is a time-involved process. So the backup vaults work very successfully because the backups will not be touched at all. However, in other instances, many organizations had backups, but the backups were affected as well. They had to recreate it and could not recover from backups.

Another challenge is that there are many mature organizations that have playbooks and backup plans in place and all governance established. However, even if all these policies and procedures are created, if there are no dry runs to practice end-to-end and tested, it may be all for nothing. The organizations must make sure that they are also used to the situation, as even if the organization has all the procedures and policies in place if they are not enacted in the middle of the crisis, they will not be effective. This might be due to people taking up more responsibilities as they perceive they are the best ones to manage it, but in a large ransomware incident, you do not have the time to discuss who's responsibility is what. The organization must be a well-oiled machine at that point. There must be good orchestration of managing all the information workflows, managing all the cyber response, and managing governance. This is a big challenge seen in organizations. 50% of the backup and recovery strategy must be the practicing of it. This must include the higher management in order to train them on how they should respond. That is the challenge. This is more from a people than a process point of view.

Another challenge is that the new emergent technology landscape, with all the new technologies such as generative AI, etc., are being embedded in organizational processes. Years ago, the technology landscape was maybe a few puzzle pieces. Nowadays it is the same puzzle, but it has way more pieces. Then it becomes critical to manage, secure and deal with the risks all these technologies create. Many organizations thus try to trim it down again to have fewer applications and systems that are also more harmonized in place. If there are five tools in place, but perhaps there is another vendor who integrates all the tools, then if the integrated tool is selected, that is four less to worry about. Organizations also might introduce group policy plans where users can only view specific items or are barred to only be able to do specific actions and cannot use third-party libraries, etc.

The primary thing that comes to mind as a critical lesson to organizations is that organizations must also involve and activate the teams and processes at the right moment, such as forensics. Some organizations sometimes are even so focused on the reactive part to reach business as usual through recovery that they forget the long-term consequences. For example, if forensics is not included at the right time prior to recovery, the forensic evidence will get deleted. In the long term, this will create problems during litigation as they will not be able to tell what happened. That is then a legal issue that must be remediated. Processes and teams such as forensics must be activated at the right time. It is not necessarily a technical issue but a more legal issue. When the ransomware attacks are analyzed, it could be seen that only about

30% of them are the technical recovery of it and remediation, but then there are the other 70% percent about the fact that you have a data breach going on. Then it is more of a legal issue and maybe a communications issue.