UNIVERSITY OF TWENTE.

Faculty of Electrical Engineering,
Mathematics & Computer Science

Proposing and Deployment of Attractive Azure AD Honeypot With Varying Security Measures To Evaluate Their Performance Against Real Attacks

Atif Mushtaq Khan

M.Sc. Thesis
Mar 2021

Supervisors:
prof. dr. ir. Aiko Pras
dr. Jair Santanna
dr. L.O. Bonino da Silva Santos

Design and Analysis of Communication Systems
Faculty of Electrical Engineering,
Mathematics and Computer Science
University of Twente
P.O. Box 217
7500 AE Enschede
The Netherlands

Preface

This thesis marks the end of my Masters program in Electrical Engineering at the University of Twente. I have conducted this thesis at the Design and Analysis of Communication Systems (DACS) research group, which is a part of the Electrical Engineering, Mathematics and Computer Science (EEMCS) department. This thesis and the work within this thesis is special to me, as it involves something which I'm passionate about; the security of the cloud services and applications.

I would like to express my gratitude to my supervisors for guiding me through the thesis and helping me in every possible way. I especially want to thank dr Jair Santanna for sparing the time for our weekly meetings and guiding me through the ups and downs during the research. It was a great pleasure to research with him and he helped me in planning and executing the work required for this thesis.

At last but not least, I want to thank my parents and my siblings for believing in me and giving me the moral support that I needed to finish the work.

Abstract

The popularity of Azure Active Directory (Azure AD) which is cloud-based Identity and access management (IAM) solution by Microsoft, has been increasing among the companies [1]. Azure AD provides the companies an affordable and easy-to-use service.It can be used as an identity provider for various first and third party applications and to manage the access privileges of the users in an organisation. The widespread use of the Azure AD by organizations for identity and access management makes it quite lucrative for the attackers to attack and gain unlawful access to the resources. In addition its innate nature of being a cloud service makes it vulnerable to security and privacy breaches linked to the cloud.

The honeypots are systems used to mimic the real system and deceive the attacker into believing that they are real systems. Honeypots are used to assist, detect and analyze attacks done on them. This is done to provide forensic information about the security breaches which can be used to provide the information about the attacks conducted on the system and how they can be prevented.

For this master's thesis, we intend to expand the application of these honeypots to Azure AD.To our knowledge this is the first time honeypots have been used with the Azure AD or cloud based IAM solutions. The honeypot is used to get the attackers to interact with the set-up and see the presence of the real-world threats that loom over the Azure AD. To achieve this goal we deployed an attractive honeypot system with various security measures depicting and representing real-world scenarios.

During the thesis, we first established a set of criteria based on the previous researches that define the attractiveness of the honeypot. The proposed planned honeypot system is then evaluated for its attractiveness against those criteria. Using that knowledge we deployed a set of 3 different honeypots with varying security hardening measures to detect the presence of real-world threats. The security measures are chosen based on how the organizations usually configure their Azure AD. The credentials for each of the set-up were leaked for one week each.

The analysis revealed the presence of the real-world threats experienced by the organizations, further verifying the attractiveness of the honeypot system. Finally, we compared the honeypots with varying security measures for their effectiveness against the detected threats. This provides us the valuable knowledge of how effective the security measures are against them. It was found that the MFA performed the best and was able to prevent the attacks. The default settings performed the worst and having custom security measures in place was able to perform substantially better than the default settings. We were also able to profile the attackers that inter- acted with the honeypot set-ups and how they interact with the set-up. Additionally, we were also able to point out some of the security flaws and shortcomings in the Azure AD and which remain an easy entry point for malicious users.

The thesis helps in establishing the foundation stone for the usage of honeypots in the IAM solutions like Azure AD and pave way for the future researches.

Table of Contents

List of Figures

Chapter 1

Introduction

This chapter explains the motivation and the goals of this thesis. Furthermore related work and the structure of the thesis are also included in this chapter.

1.1 Motivation

Over the last decade, the increasing utilization of the Cloud and its applications has made the Azure Active Directory (Azure AD) quite popular among organizations. Azure AD is used to manage more than 1.2 billion identities and access privileges. Each day around 8 billion authentications are made using the Azure AD [1]. Azure AD helps organizations who want to build their own applications, by providing world-class identity and access management services. It provides the organizations a low-cost, reliable, and easy-to-use way to enable single sign-on for thousands of first and third-party applications like Office 365, Salesforce.com, and others. In addition to features like single sign-on, Azure AD provides reliable access management, multi-factor authentication, and usage monitoring. It also provides detailed logs for security and auditing purposes.

The popularity of Azure AD has bought it into the radar of the attackers who want to gain access to the Cloud resources to satisfy their malicious intentions. As per Microsoft nearly 0.5% (1.2 million), of Azure AD accounts are compromised every month, due to lack of Multi-factor authentication [2]. Another study also conducted by Microsoft revealed that due to password reuse (and other factors), more than 44 million Azure AD accounts have been Compromised recently [3]. Password Spray [4], a type of attack in which multiple accounts are attacked with the help of commonly used password, tend to be quite effective against Azure AD.

To counteract these attacks the security of the Azure AD can be hardened using certain security measures. However many organizations are not able to completely harden the security posture of the Azure AD and leave room for these attacks. Weaker security measures or misconfiguration of the access policies is often the cause of these breaches. The high number of attacks along with the substantial success rate has made it of prime importance to focus on the security of the Azure AD accounts.

To protect against such attacks and provide valuable information about them, honeypot systems could be used. The Founder of Honeypot project Spitzer (2002) [5] described honeypots as " honeypot is security resource whose value lies in being probed, attacked, or compromised" or "an information system resource whose value lies in unauthorized or illicit use of that resource" Spitzner (2003) [6]. Honey-pots are systems that try to mimic the real systems and try to fool the attackers. During the interaction of the attacker with the honeypot system valuable information about the interaction is being logged and stored. This information can be used to learn about the attacks and the effectiveness of the various hardening measures against those attacks.

1.2 Related Work

In this section, the related work and research are evaluated and discussed. Honeypots have been used in past researches to analyze, detect, and devise mitigation measures against real attacks. Some of the related researches are as under:

Liston (2002) [7] created a honeypot program known as LaBrea. It uses the unused IP of the network and answers to the malicious connection attempts. It takes a long time to answer these connection requests and hence wastes the time of the attackers.

Krawetz (2004) created [8] fake spam email sending tool as a honeypot to gather information about the spammers. This information was used to find the details about the spammers who use open proxy relays to stay anonymous while conducting the attack.

Virvilis, Serrano, Vanautgaerden (2014) [9], was able to deceive the attackers and protect the valuable accounts and information inside the network. To do that they created fake files with misleading names like "secret" "important" or "confidential" on their network devices. In addition to this, they also created fake domain name system records and HTML content on the network for this purpose.

Akkaya Thalgott (2010) [10], provides valuable information about the legality of the honeypots. It answers how the network administrator can obtain the information about the attacker legally and what laws restrict the usage of honeypots in the United States and European Union.

Middleware (2019) [11] used the high interaction honeypot to lure the attackers and collect the information about the attackers. This information about the different kinds of attacks conducted through RDP after acquiring the credentials was used to analyze the different attacks been conducted.

Brown et al (2012) [12] were among the first to use the honeypots for the security of cloud infrastructure. The authors deployed low to medium interaction honeypots on Cloud infrastructure of Amazon EC2, Microsoft Azure, IBM smart Clouds, and Elastic Hosts to analyze the incoming traffic and the interactions with the honeypots. They covered various regions across the world by deploying 42 instances and analyzed the malicious traffic that interacted with them.

Davide Bove, T. Muller (2018) [13] also used the public Cloud infrastructure of Google, Amazon, and Microsoft to deploy the honeypot systems on the Cloud. They set up the honey-pots to mimic the SSH and VNC services and were able to collect the information about the attackers based on the logs they generated while interacting with the honeypots. The honeypots deployed were mainly low interaction honeypots and were designed to allow attackers to access only specified types of service only.

While a lot of research has been conducted on the usage of honeypots for various applications. All these papers and researches mostly focus on the traditional computing resources or services like RDP, SSH, VNC, etc. To our knowledge, no research has been conducted on the deployment of honeypot for identity and management services like Azure AD. However, they do provide a detailed insight into how the honeypots need to be deployed attractively and how to make them accessible to the attackers who are intended to interact with them.

The above researches even though are not about the identity and access management services (Azure AD), still provide significant information in key aspects of honeypot deployment. They give detailed information about the ways the honeypot can be used, and set-up. How the honeypot system as a whole can be deployed to make it more attractive and how we can make it look more realistic and reachable to the attackers. How the credentials can be leaked on the internet to make them available to the attacker.

All this knowledge forms the ground basis for our research in this thesis and using this knowledge we will be able to deploy an Azure AD honeypot system which will be used to analyze the attacks and evaluate the performance of security measures against them.

1.3 Goals & Contribution

This thesis aims to evaluate the security hardening measures of the Azure AD when the credentials of the users get leaked. To have this knowledge we have to first establish a framework for the honeypot system which is attractive, legal, and non detectable to the attackers. After that three versions of the honeypot systems are deployed with varying security postures. The deployed honeypots are then evaluated for their performance based on their effectiveness against the identified attack vectors.

When the attacker interacts with the honeypot system, the logs generated from their interactions are recorded and analyzed. Additionally, we tried to identify the tools and other details about the attacker.

The contributions in this research are as follows:

- Establishment and evaluation of the criteria for an attractive Azure active directory honeypot system. This provides a framework for designing and evaluating a honeypot system before it has been deployed.
- Deployment of the honeypot system and leaking the credentials to identify the presence of the real-world attacks. This helps in visualizing the real threats that loom over the real systems in the wild and also verify that the attractiveness of the deployed honeypot system.
- Information from the data logs analyzed for the evaluation and comparison of the security measures against these attacks. This provides valuable information about the effectiveness of the measures against the threats.

The results of the analysis and the consequent knowledge may help in choosing the security measure against the attacks and provide an insight into how much better they are in comparison to each other.

1.4 Research Goals

The lack of substantial research regarding the use of honeypots in Azure AD marks the ground basis for our research. During our research, we will answer how to set up an attractive honeypot system and how the various security measures in the Azure AD perform against real-world attacks.

During the research we want to extend the usage of honeypots to Azure AD and intend to answer the below questions:

- What criteria define an attractive azure AD honeypot set-up and how they are met.
- What different honeypot set-ups are deployed and what real-world attacks have been observed.
- How the various security hardening measures perform against these attacks observed.

1.5 Structure

The structure of this thesis is as follows:

- In chapter 2 the necessary background knowledge about cloud computing, honeypots, Azure AD, and the Enterprise apps deployed is provided. Furthermore, the criteria for setting an attractive honeypot are discussed.At the end of the chapter the proposed honeypot system is evaluated against those criteria.
- The subsequent chapter 3 describes the deployment of the honeypot set-up and how the credentials were leaked online. It further briefs on how the logs generated are captured and retrieved for analysis of the attacks.It also includes the problems faced while deploying the honeypot system and a preliminary analysis of the data collected.
- Chapter 4 provides a detailed analysis of the attacks conducted during the entire up-time of the Azure AD honeypot system. A comparison between the performances of the security measures is discussed.
- Final chapter presents the conclusion of the thesis and what future improvements and research can be conducted.

Chapter 2

Background and Criteria for an Attractive Honeypot Set-up

This chapter provides the necessary background information on the various systems, applications, and techniques used during the thesis. The chapter highlights the core difference of cloud services from regular infrastructure. Additionally, information about the Azure AD, Single Sign-On, and Enterprise apps deployed in this thesis are also given. Furthermore, the concept of honeypots is introduced, which is the backbone of the experiment.

This chapter will mainly focus on the requirements and criteria for an attractive honeypot setup. For this thesis, an attractive honeypot is defined as a setup that can fool an attacker into believing it to be a real system. The chapter also includes a section wherein we discuss various resources that will be deployed to make the honeypot system more attractive. At the end of the chapter, we will analyze whether the criteria mentioned at the beginning of the chapter are met while establishing the honeypot set-up.

2.1 Background Knowledge

This section will give a brief overview of the various technologies related to this research.

2.1.1 Honeypots

Honeypots [5] are the systems having no monetary value as such for any business. They are placed in or instead of the real systems for various security reasons. Anyone interacting with these systems is considered suspicious and one with malicious intent. Honeypots are used to waste the time of attackers by making them interact with the dummy system instead of harming the real network/systems. They are also used to gather and analyze the information about the attacker and the various techniques used while interacting with them. They can gather all the interactions of the attackers in the form of readable and analyzable logs, thereby helping in patching and securing the vulnerabilities of the system [14].

Numerous types of honeypots [15] are distinguishable according to their capabilities, necessities, and results. However, based on the capabilities of the honeypot to mimic the real system we have four different types of honeypot systems:

- Low-Interaction Honeypot Systems: These Honeypots are very limited in nature and can run few services. These services provided are also very constrained compared to the actual services. Only a small amount of information is obtained from these systems as the interaction between the attacker and these systems is very limited. However, utilizing these sorts of honeypots incredibly decreases the chance of system abuse by the attacker.

- Medium-Interaction Honeypot Systems: These types of honeypots are more detailed in structure and deployment than the low-interaction honeypots. They offer more interaction to the attacker and hence can extract more information from the attackers. The attackers try and spend more effort and time interacting with them. They try complex techniques and tools to gather and access more resources, which reveals their plans and is quite useful for the security admins. However, the more complex nature of these honeypots poses a risk of the abuse of set-up by the attacker.

- High Interaction Honeypot Systems: These real systems running real services. Interaction between the attackers and the systems is detailed and provides broad information about the attacks. Most of the attacks could be analyzed using these honeypot systems. However there are certain drawbacks of using these systems due to the high level of interaction there is a serious risk that the attacker breaks the honeypot setup and gains access to the actual assets, hence proper isolation needs to be put in place to prevent this threat. These systems are costlier and demand more time for their deployment.
  In our thesis, we are using this type of honeypot, as it is an actual and real Azure AD system.

2.1.2 Cloud Services

Cloud Computing is defined as the availability of on-demand scalable and virtualized resources over the Cloud (Internet) [16]. Cloud computing helps in running the infrastructure at a low cost and efficiently by scaling as per the business demands. Public Cloud providers have a really simple business model, as a client you need to spend only for the resources you utilize which terminates the necessity for cloud computing users to plan far ahead. Armbrust et al (2010) [17]. Some of the commonly known examples of public cloud provides are Microsoft Azure, Amazon Web Services, Google Cloud Platform, and IBM Cloud. Public Cloud Providers which provide access to cloud infrastructure, resources, and administration to common individuals in contrast with private data centers for companies, are the main focus of this thesis.

As per the National Institute of Standards and Technology (NIST) three different models of cloud computing are:(Mell and Grance 2010) [18]:

- Software as a Service (SAAS): These are applications that run natively on the cloud and can be accessed through their web interface or via the API. All the underlying details of network infrastructure are hidden from the End-users. The infrastructure is operated and maintained by the service provider.
- Platform as Service (PAAS): In this type of system the applications are deployed and run by the customer on the infrastructure provided by the cloud service provider. In other words, the infrastructure is operated by the provider along with assets, operating system, and hardware but the client can deploy and run his application on this platform.
- Infrastructure as a Service (IAAS): In this system, the provider provides the user with the infrastructure and the end-user is given full control over the operating system, applications, storage, and network configurations. Only complex and hardware-related tasks related to the hardware are with the provider of the service.

The security of cloud infrastructures is by itself an interesting subject, however, the security of the identity and access management system like Azure AD is the main focal point for this thesis.

2.1.3 Azure Active Directory (Azure AD)

Azure Active Directory [1] is a multitenant Identity and access management service by Microsoft. Azure AD is Cloud-based and assists users belonging to an organization to sign-in and authenticate various resources. It is widely being used by the organization for its ease of use and reliable operation while reducing the overall cost.

- Azure Active Directory can be used to access enterprise applications, Microsoft 365 office suite, and Azure Portal. It can be also used as an identity provider for SSO for other third-party applications as well

- Azure Ad can be incorporated and linked to the internal directory of the corporations. This way it can be used to control access to the internal applications running on the intranet.

It helps the organizations by streamlining the method of identity management services. In addition to being cheaper, scalable, and reliable, it also increases the overall security posture. Azure AD logs are highly detailed and provide a great insight into any interaction happening which offers great assistance in monitoring the access to the cloud resources and applications. This is a great way of securing and monitoring access to the application and other resources.
Refer to the appendix for all the information contained in the logs of the Azure AD

2.1.4 Single Sign-On (SSO)

Single sign-on (SSO) [19] is a scheme by which a user can authenticate multiple applications, systems, and services using one set of credentials. SSO is being used by various organizations to make the job of managing the credentials easier. Azure AD as an identity provider can be used with several enterprise applications for single sign-on. This makes it easy and hassle-free for both the system administrators and the users of those applications.

Single Sign-on improve the productivity and security of the organizations due to the below reasons:

- Due to the single sign-on the users have to remember 1 sets of credentials, hence it reduces the password fatigue and inhibits the users from using weaker passwords or sequential passwords for different services/ applications.

- It reduces the time wasted by the users in re-entering the same or different credentials for different applications.

- It reduces the cost and time spent by the IT department in troubleshooting the issues involving the passwords.

2.1.5 Ethical and legal and other issues related to honeypot

Restricting the abuse of the system can be accomplished by constraining the interaction of the attacker with the honeypot system. This however will reveal the reality of the setup that it is a honeypot instead of a genuine system. Ethical obligations play a major role in this regard. To fulfill both the prerequisites a balance between the two processes permitting higher interaction and maintaining a strategic distance from the abuse needs to be achieved.

Ethically set up should be such that it should not be utilized to hurt other systems by the attacker. With regards to the legality, the privacy laws of the nation in which the setup is established are applied. Privacy-related issues concerning honeypots are quite evident. The data collected by the honeypot system contains information like IP address, time of interaction, and geo-location of the attacker. All this information is considered as personal information as per the EU cyber laws [20] and special care needs to be taken while collecting and using this data.

Collecting data with the intent of avoiding such attacks in the future justifies the purpose and allows the operator of the honeypot to do so. However special care needs to be taken before the data is made public or the data set is published. Sokol, Misek, and Husak (2017) [21], recommend the results must be anonymised, as the results contain personal information that is protected under cyber laws like GDPR. However one can argue that there is no conclusive answer to whether it is legal to collect this data containing the personal information of the attacker or not. It boils down to various factors like :

- What is the reason for collecting the data from the honeypot set-up
- How the data will be processed to protect the privacy of the attacker.

2.2 Criteria set for the attractiveness of the honeypot

One of the main research questions is how to set up an attractive honeypot setup. This question is of prime importance because the attractiveness of the set-up is the main thing that brings the attackers towards the set-up and makes them interact with it.

Based on the previous researches we have devised a set of criteria that dictate whether the honeypot system is deemed attractive or not. This helps us in creating a framework for deploying the honeypot system in the later chapters successfully.

The 3 main criteria for a successful honeypot system are:

- Whether the attackers after visiting the system finds out that he is interacting with the honeypot set-up instead of the real system [22]. The honeypot should be deployed in such a manner that it makes the attacker believe that he is interacting with the real system. By doing so, the attacker will end up spending more time interacting with the system to have some unlawful gains.
- Whether the attacker wants to interact with the honeypot set-up. Interacting with the set-up requires time and effort from the attacker [23]. So to make the attacker interact with the honeypot, there should be something that will benefit the attacker.
- Whether the honeypot set-up is made secure enough that the attacker won't be able to conduct malicious activities on other systems using the honeypot [24]. Also from the ethical and legal point of view care should be taken that no personal information regarding the attacker is revealed during research [20].

If the honeypot system can meet all the above requirements, it will be regarded as an attractive honeypot set-up and will resemble the real-world system closely. This will try to answer us the research question 2 and 3 properly and more effectively.

2.3 Resources to make honeypot set-up attractive

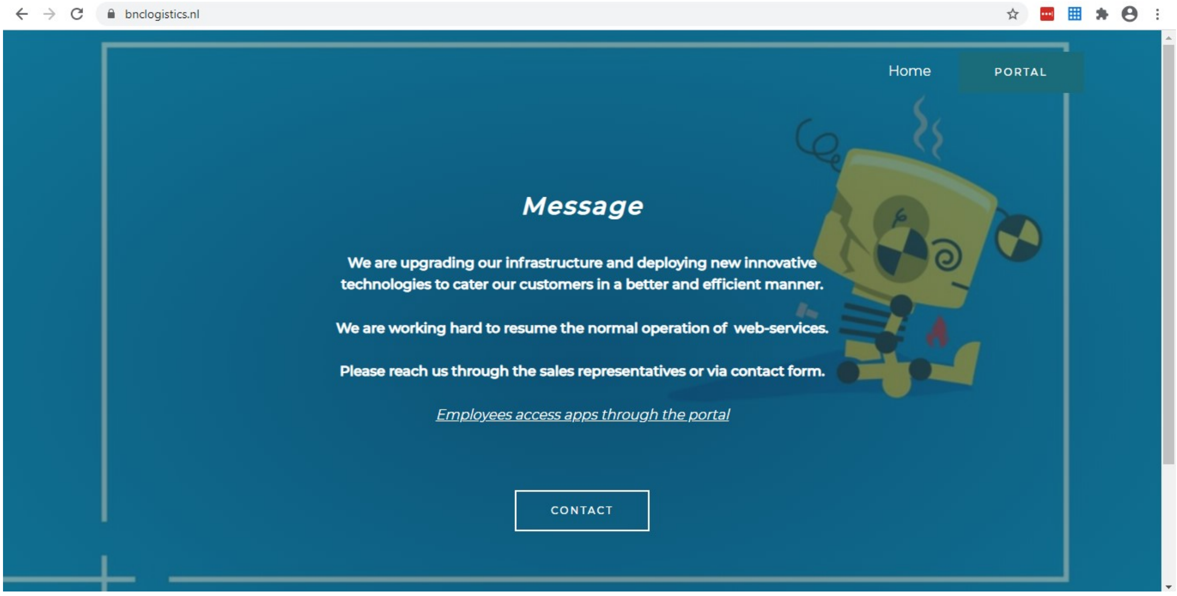
The main purpose of the honeypot system is to gather information regarding the attacks conducted by the attacker. To do that successfully the honeypot system should be attractive to lure the attackers towards it. The attacker should not be suspicious that the system he is interacting with is fake and not the real one, at the same time the system should be isolated and protected enough so that he can't gain unwanted access by lateral movement.

In the thesis, a high interaction honeypot system has been considered. A real Azure Active Directory has been established using the domain name registered under the name of the fake company called BNC-logistics. The users/employees in the directory are setups with fake Dutch names, fake contact info, and different job description referring to various job positions. This is done to make it believable that the users in the Azure AD are real and make it look more realistic to the attacker

Some of the main components of the honeypot system that will make it more attractive are:

2.3.1 Website

To increase the attractiveness of the honeypot set up, a website was hosted on the domain registered under the fake company's name (bnclogistics.nl). The website states that the company is upgrading its infrastructure and because of that normal operations are affected. The statement about upgrading was added to cover the fact that the honeypot is established few weeks prior and also the domain is recently registered. This makes the attacker doubt less about the authenticity of the honeypot setup.



Message

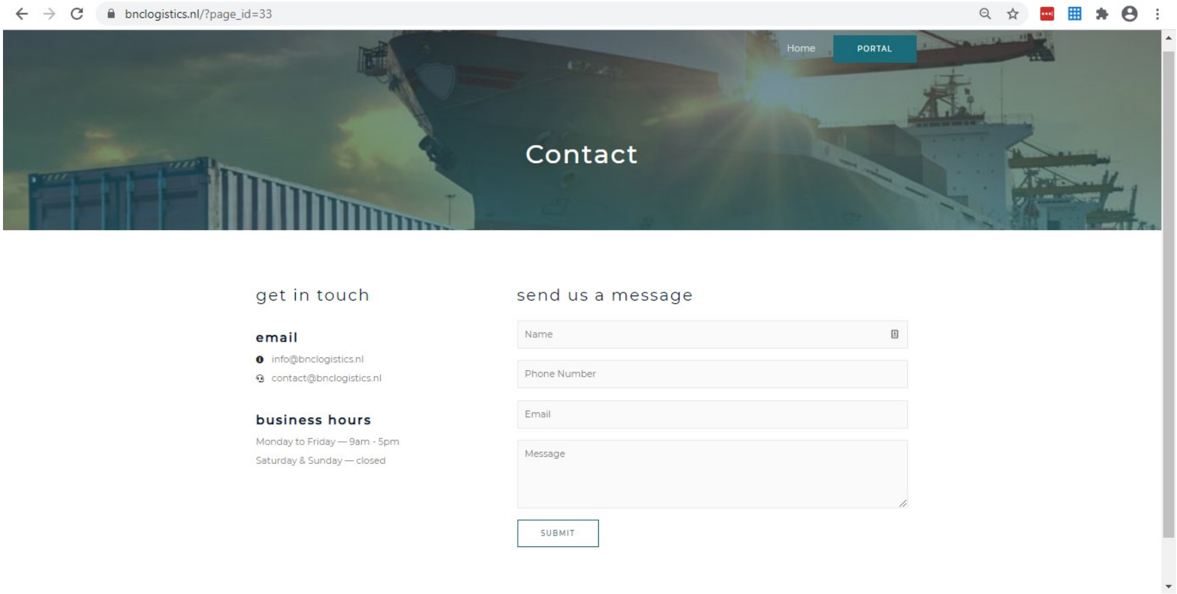
We are upgrading our infrastructure and deploying new innovative technologies to cater our customers in a better and efficient manner.

We are working hard to resume the normal operation of web-services.

Please reach us through the sales representatives or via contact form.

Employees access apps through the portal

Figure 2.1: Homepage of the website (BNC-Logistics)



get in touch

email

info@bnclogistics.nl
contact@bnclogistics.nl

business hours

Monday to Friday — 9am - 5pm
Saturday & Sunday — closed

send us a message

Name

Phone Number

Email

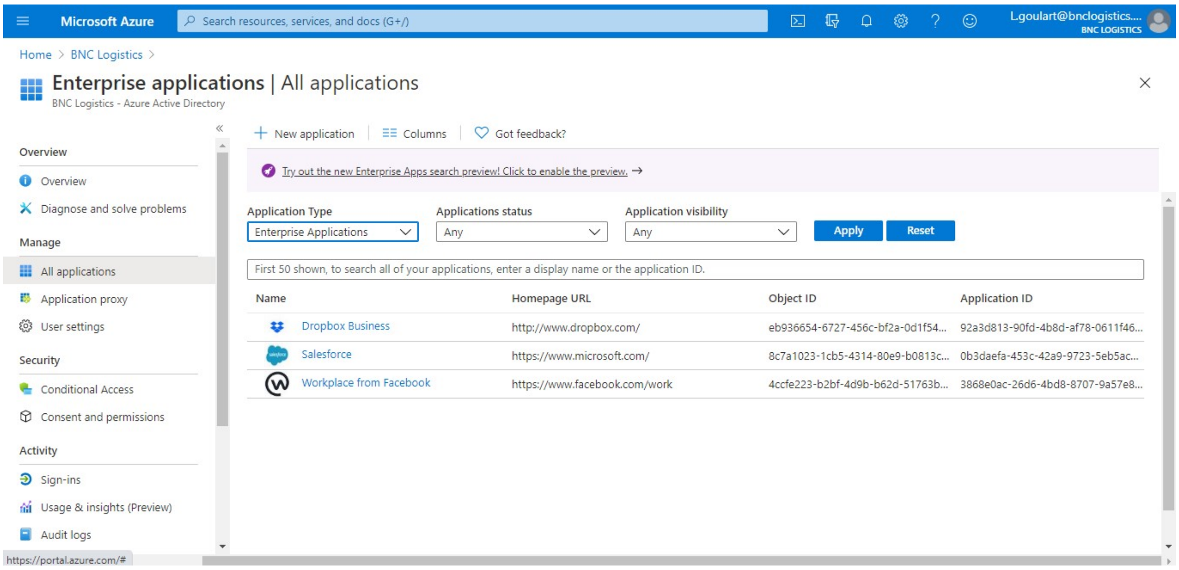
Message

Figure 2.2: Contact Page on the Website (BNC-Logistics)

To make the website look a little bit more realistic, a contact page and contact form were also created, so that the customers can contact in case of urgent queries. In addition to all this, a button and a note addressing the employees to log in and access the applications via the portal were added to the homepage. The button was added so that the attacker after visiting the website will be able to directly reach the login portal for Azure AD honeypot setup.

2.3.2 Azure Enterpise Applicatiions

Azure AD can be used to manage the access to third-party enterprise applications as well as applications running on premise [25]. Azure AD can be used as an identity provider for a single sign-on service and can store all the identities required for authentication in different applications.



Enterprise applications | All applications

BNC Logistics - Azure Active Directory

Try out the new Enterprise Apps search preview! Click to enable the preview. →

Application Type: Enterprise Applications
Applications status: Any
Application visibility: Any

First 50 shown, to search all of your applications, enter a display name or the application ID.

Name	Homepage URL	Object ID	Application ID
Dropbox Business	http://www.dropbox.com/	eb936654-6727-456c-bf2a-0d1f54...	92a3d813-90fd-4b8d-af78-0611f46...
Salesforce	https://www.microsoft.com/	8c7a1023-1cb5-4314-80e9-b0813c...	0b3daefa-453c-42a9-9723-5eb5ac...
Workplace from Facebook	https://www.facebook.com/work	4ccfe223-b2bf-4d9b-b62d-51763b...	3868e0ac-26d6-4bd8-8707-9a57e8...

Figure 2.3: Enterprise applications in the Azure AD

The enterprise apps are deployed to make the honeypot system attractive to the attacker and make him interact with the system to gain access to them. We are planning to deploy the apps in such a way that to have the access to the applications the attackers have to elevate their account privileges/access or make a lateral movement to some other account that has access to this application.

The three enterprise applications that we chose for the setup and makes sense for a logistics company are as follows:

- **SalesForce** [26] is one of the leading CRM (Customer Relationship Management) running on the cloud as SaaS. It has more than 800 features to support

various jobs like new leads, sales, and closing deals.
Salesforce in our Azure AD set-up was selected, as it makes sense for the logistics company where they can manage the customers and sales details using this application. While deploying the application, the SSO feature was enabled and the login page of the application was set-up in such a way that it accepted only the authentication via Azure AD. The company's webpage was also displayed on the left half of the login page. All this was done to give it a more real and professional feel.
The application aimed to act as an asset for the attacker, such that he tries to gain access to the application and eventually to the sensitive private customer or sales data.

- **Workplace** [27] is a mobile or web app developed by Facebook, running on the cloud, and is used to communicate with the team members. The service provides features like group chats, voice and video calls along with access to social media events, live video tools, and profiles.
  This application is chosen as in the case of any company the employees need to communicate with each other and with clients or suppliers. This application is quite famous and most of the people/attackers are familiar with the workplace and know what it is used for.
  This application is deployed to attract the attacker. The attacker may be interested in knowing the internal communication between the employees of the organization or with clients/suppliers, which may leak trade secrets and confidential information. All this increases the attractiveness of the honeypot system.
  The application is visible to the attacker and since the SSO is enabled for this application the attacker will try to access the application by gaining the access to accounts which has this application access.

- **Dropbox Business** [28] is a file-sharing application, that is used by companies and organizations. It is used to share the files within the organizations and can be also used to store the files securely on the cloud for easy access from anywhere. Dropbox can be also used to share the files between the employees or with the clients/suppliers.
  An application like Dropbox is attractive to the attackers, as it can give them access to the files stored and shared between the users. The attacker tries to gain access to the application and interact with the setup in doing so. This is the reason why this application was added to the Azure AD honeypot system.

These applications make the system look realistic and gave a fake sense of reward to the attacker, in case he manages to get access to them. All these applications help attract the attacker so that he can spend time and conduct the attacks on the honeypot system.

2.4 Whether the goals set to analyze the attractiveness of honeypot are met

Analyzing the honeypot set-up against the criteria set in section 2.1, we can argue whether the proposed honeypot system is attractive, safe, and legal or not. The below are some of the key justifications about the attractiveness of the proposed honeypot system.

- The Azure AD apps, using the names and job profiles which make sense for the Dutch logistics company, creating a website, registering a domain name matching the name of the company makes the whole honeypot setup look realistic.
- The whole set-up is made to look like a real system where resources (apps like Salesforce, Workplace, and Dropbox) are available to the attacker, to interact with them. Hence it can be argued that the deployed set-up is attractive to the attacker based on the specified criterion that there is something for the attacker to get benefit from.
- By restricting the access privileges of the compromised accounts we made sure that the attacker won't be able to harm the honeypot system or use it for carrying the malicious activities. The whole set-up was restored periodically to make sure that the attacker isn't able to use it for illegal activities in case he finds a loophole in the security of the set-up.
  It is also made sure that the personal information of the attacker during the analysis phase of the research is made anonymous so that we are in inline with the ethical and legal considerations.

From the above, we conclude that the required criteria are met successfully and the honeypot system is an attractive setup.

In this chapter, we were able to provide the background information related to the thesis. Furthermore, we defined the framework for an attractive honeypot setup and what resources are required to meet those criteria. We concluded the chapter by evaluating whether and how those goals are met.

Chapter 3

Defining Different AAD Honeypots

The chapter provides information about the honeypot architecture, how it has been deployed, and how the attackers are lured towards it. The various expected and observed threats against the Azure AD in the real world, various types of credentials, and security postures are also highlighted in this chapter. The later part of the chapter deals with the three different honeypots deployed and the reason for their variations. At last, we have a section dealing with the retrieval of logs and preliminary detection of the threats. Additionally, we also have a section mentioning the roadblocks we faced during the deployment of the honeypot.

3.1 Different types of expected threats

As per the definition of Howard and Longstaff (1998) [29], attacks are defined as "a series of steps taken by an attacker to achieve an unauthorized result". This is quite important as the honeypot will record several logs and records which can be grouped as a single incident/attack, as they are used to achieve one final distinct outcome. By doing this the overall amount of data that need to be processed is greatly reduced. It also provides a better overview of the attacks and how much similar they are to real-world threats.

To identify the real-world threats we consider the MITRE ATT&K matrices. MITRE ATT&K [30] is a knowledge base of real-world attacks, tactics, and the techniques used by the attackers all around the world The attacks in the MITRE ATT&K matrices are what the real system experience in the real world. Below are the threats that are considered for the Azure AD honeypot:

- **Create Accounts/Persistent Attack** In this form of attack the attacker may invite the guest user to the directory. In case the attacker manages to access the account with administrative rights, he can create users and will end up having persistent access to the directory.

- **Account manipulation/ Lateral Movement** In this form of attack the attacker can manipulate the access that he has for his account, or he can try to get access to the accounts that have administrative rights or elevated access. This movement is known as lateral movement. By doing so the goal of the attacker is to get access to accounts which has access to most of the cloud services/resources and apps.

- **Brute Force** In this form of attack the attacker tries to brute force into the accounts using commonly used passwords. The brute force is quite an important and prevalent form of threat against the Azure AD. usage of the same password across the different account, common password string, weak password policy, etc are the main reason for the success of brute force in getting the unknown password

- **Password Spray** In this form of attack the attacker instead of brute forcing a single user, sprays the entire user list or certain chosen users with a password or list of passwords. The attacker first tries one password for all the users before moving to the next password. This method is quite effective in case the users have a weak password and no MFA enabled.
  From the detection point of view, this type of attack is really difficult to distinguish from an isolated failed login. Sometimes the attacker even waits between the two passwords to prevent the account from being locked out. This further makes the detection of this attack a difficult and tedious job.

- **Vandalism** This attack is related to the deletion of the accounts, files or tampering with the settings of the honeypot system. The goal of this type of attack is more destructive than disruptive. In our honeypot system we have put counter measures against this type of attacks and the probability of attacker being able to conduct vandalism is very low.

- **DOS attack due to Lockouts** This is a special kind of attack as the attacker can do it voluntarily or involuntarily.
  Voluntarily is when the attacker wants to disrupt the access of the user to the Azure portal. He can do that by deliberately making several consecutive failed sign-in attempts resulting in the lockout. Once the account has been lockout the access to that can be restored only after the completion of the period for which the account was locked. Unfortunately, there is no other way to restore access to the locked account.
  Involuntarily DOS attack happens when the user tries to brute force into the

account or password spray the users. The attacker's only goal during these processes is to have the access to the accounts which are under attack, but while doing so he may cause account lockouts, which renders those accounts inaccessible to even legitimate users for a particular period.

3.2 Personalized Honeypot Content

The content in the honeypot system is what makes it useful or useless when it comes to collecting information about the attacks. The content of the honeypot system should match the nature and description of the honeypot setup. The content should be such that the attacker finds it interesting and fruitful to spend his time and resources while interacting with the setup. The assets that the attacker can gain access to, are what makes the honeypot attractive to the attacker. The attacker tries to gain access to the assets which can be enterprise apps/services or cloud resources. The attacker's lateral movement while trying to gain more access to the assets can be logged, analyzed, and studied to learn the behavior of the attacker.

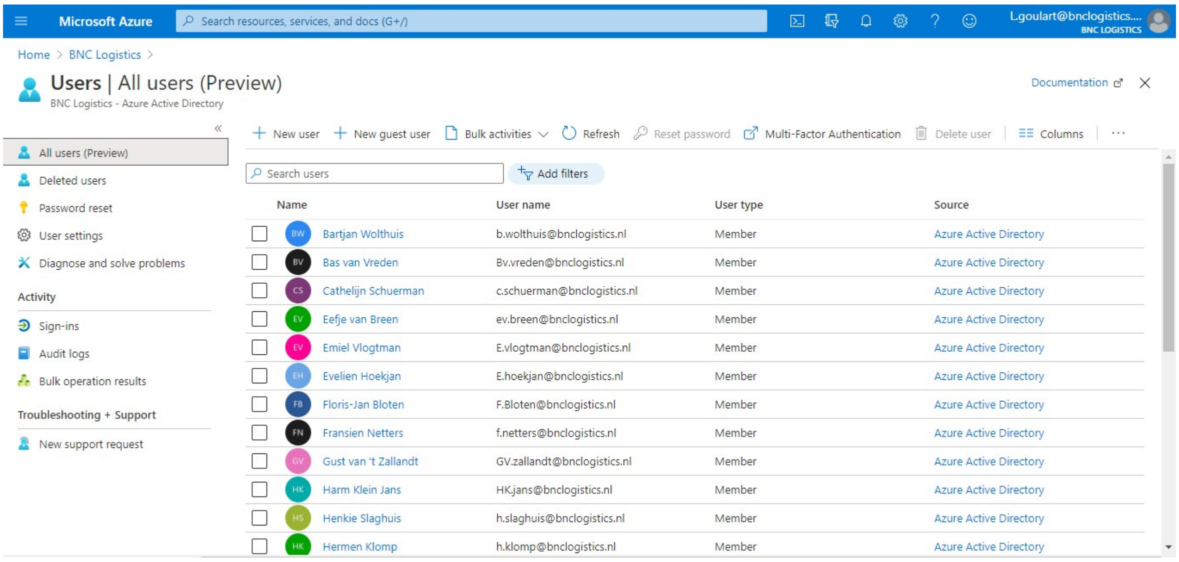
In the thesis, to have a meaningful and attractive asset to attract the attacker, three enterprise apps are being deployed namely Salesforce, Workplace by Facebook, and Dropbox Business. The access to these apps is managed by the Azure AD using SSO and provisioning. The Azure AD comprises 56 users, which are generated using random English and Dutch names. All the users are employees of the BNC-Logistics. Other details like Employee ID, Email address, and job description are also mentioned.

The total users are divided between these 3 types as follows::

Type of User	**Admin Access**	**MFA Enabled (All 3 Runs)**	**App Access**	**No. of Users**	**Credentials Leaked**
Type I	No	No	Yes	40	No
Type II	Yes	Yes	Yes	6	No
Type III	No	No	No	10	Yes

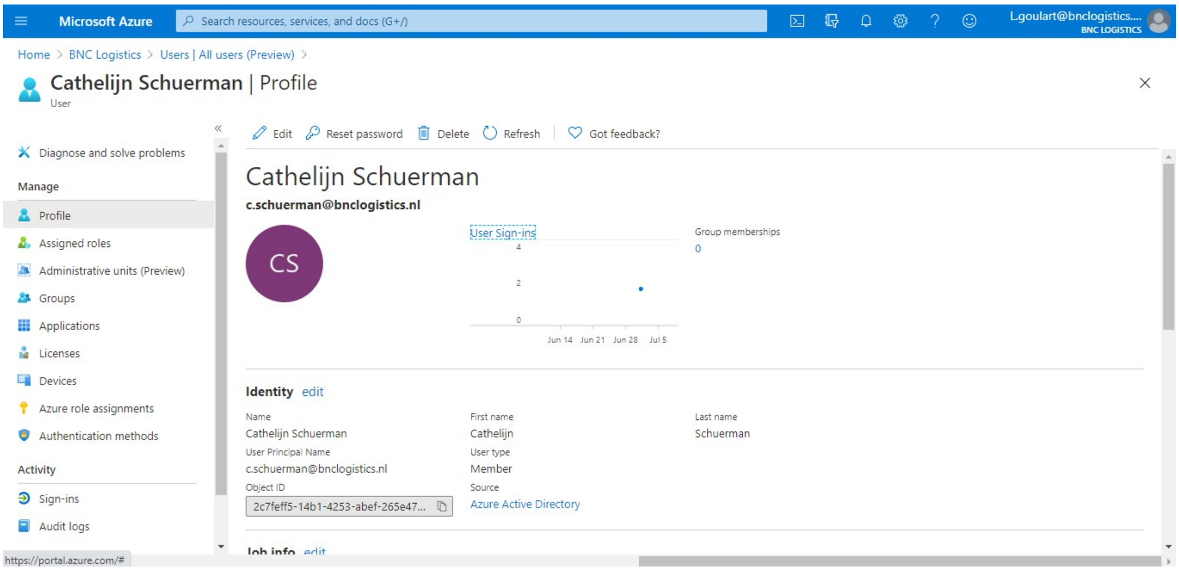
- Type I: Standard users with no administrator access, but having access to the apps, relevant to their job description.
- Type II: Users having administrative access along with access to the apps
- Type III: Users with no access to the apps and no administrative roles. The

credentials of these users will be leaked online. They act as the entry point for the attacker.



Name	User name	User type	Source
☐ BW Bartjan Wolthuis	b.wolthuis@bnclogistics.nl	Member	Azure Active Directory
☐ BV Bas van Vreden	Bv.vreden@bnclogistics.nl	Member	Azure Active Directory
☐ CS Cathelijn Schuerman	c.schuerman@bnclogistics.nl	Member	Azure Active Directory
☐ EV Eefje van Breen	ev.breen@bnclogistics.nl	Member	Azure Active Directory
☐ EV Emiel Vlogtman	E.vlogtman@bnclogistics.nl	Member	Azure Active Directory
☐ EH Evelien Hoekjan	E.hoekjan@bnclogistics.nl	Member	Azure Active Directory
☐ FB Floris-Jan Bloten	F.Bloten@bnclogistics.nl	Member	Azure Active Directory
☐ FN Fransien Netters	f.netters@bnclogistics.nl	Member	Azure Active Directory
☐ GV Gust van 't Zallandt	GV.zallandt@bnclogistics.nl	Member	Azure Active Directory
☐ HK Harm Klein Jans	HK.jans@bnclogistics.nl	Member	Azure Active Directory
☐ HS Henkie Slaghuis	h.slaghuis@bnclogistics.nl	Member	Azure Active Directory
☐ HK Hermen Klomp	h.klomp@bnclogistics.nl	Member	Azure Active Directory

Figure 3.1: Users in Azure Active Directory(Azure AD)



Cathelijn Schuerman
c.schuerman@bnclogistics.nl

Identity edit

Name: Cathelijn Schuerman
First name: Cathelijn
Last name: Schuerman
User Principal Name: c.schuerman@bnclogistics.nl
User type: Member
Object ID: 2c7feff5-14b1-4253-abef-265e47...
Source: Azure Active Directory

Job info edit

Figure 3.2: Details Of the User in the Azure AD

During the setup of these enterprise apps, the domain name and the email address of the fake company were used. The login URL/page of these apps also contains the company name. Salesforce login page was also modified to show the webpage of the company on the right half side of the page. All this makes the attacker believe in the validity of the company. Regarding the problem of recent creation dates of various users and apps, the story of undergoing a migration to the cloud makes it believable that the users and apps are real instead of fake ones.

The data stored in the apps act as the assets for the attacker and the attacker may try to make the lateral movement and gain access to the account which has access to the apps he is interested in. In addition to the apps, the accounts with administrative rights are also quite enticing to the attacker, as he can gain access to the cloud resources and another service if he can get access to those accounts by making lateral movement.

3.3 Various types of Credentials used

The users or the systems are authenticated with the help of information known as credentials. Credentials are in various forms like secret knowledge (username and password), certificate, token, fingerprint, etc [31]. Sometimes additional information besides this is used to authenticate the user such as time, location, or IP address. Among all the above methods the combination of password and username is the most common form of credentials for authentication purposes.

As per the survey conducted by NordPass [32], a password management software company, an average person has more than 100 username/password combinations. Keeping track of these many unique and difficult passwords is not humanely possible. Stobert and Biddle (2014) [33] conducted an experiment that proves the exact point that the user even-though has to keep track of a large number of password/username combinations, most of the passwords are being reused for different usernames. In their study, they found out that 26 out of 27 people reuse the same password for different usernames about different accounts.

As per the NIST Special Publication, 800-63B [34], "Humans, however, have only a limited ability to memorize complex, arbitrary secrets, so they often choose passwords that can be easily guessed.". The fact that humans reuse the same password for many accounts and easily guessable passwords are one of the main concern and reasons for the security breaches. A weak password can be easily guessed by the attacker by performing a password dictionary attack using commonly used passwords.

Furthermore, malicious tools such as Keylogger and credential-stealing malware are easily available on the internet and can be used to record the password and other details typed by the user and forward them to the attacker. Websites and emails involved in phishing are another common way in which credentials get stolen. In a phishing attacker, the attacker creates a fake similar-looking website where the user is led to and is made to login with his credentials. The credentials entered are easily seen by the attacker who manages that fake website/login page.

To improve the overall security of the credentials a multi-factor authentication can be used. A multi-factor authentication [35] adds additional steps to the authentication process. The user after sharing the first set of credentials is asked to share another piece of information from a device that needs to be physically present with the end-user like a smart card, fingerprint, or a text code on the user's phone. By having an additional step, the security of the accounts is increased even if the credentials used during the first step are weaker. By using Multi-factor authentication the probability of an account being compromised is very low even If the credentials are known or stolen because the device used for the second step is expected to be with the user.

However despite having a strong and positive impact on the security of the credentials a large number of corporations are not using multi-factor authentication. A survey conducted by the KnowBe4 [36], involving 2600 IT professionals revealed that about 38% of larger corporations do not use multi-factor authentication (MFA), and neither do 62% of smaller to mid-sized organizations. There are several reasons why companies don't seem to be eager in using MFA, despite its benefits. Annoyance to the users, no clear apparent and clear benefit to the management, risk of preventing intended user from logging in successfully, not 100% secure and foolproof, time-consuming, etc are some of the reasons why companies are still not going full force for the MFA.

Some even argue that it is much more convenient and efficient to use stronger password policies like specifying the password requirements, not allowing reuse of passwords, and changing passwords after a fixed interval of time than the hassle of implementing the MFA in their organisations [37].

3.4 Honeypots with varying security measures

The variations in the security posture of the honeypots are based on the fact that organizations use different security measures for their Azure AD as highlighted in the above section. Deployment of various honeypot set-ups helps us in evaluating the performance of security hardening measures against each other and see which of the security hardening measure performs better against the real-world attacks detected.

The different honeypot set-ups that were deployed are as below:

3.4.1 Honeypot with default security measures

In this set-up, the default settings for the Azure AD account were used. This set-up was deployed to replicate the organization where no special security measures are put in place. Some of the key security drawbacks in this set-up are:

- The users can invite the guest users.
- The users once logged in the portal can see all the users and the applications registered in that directory.
- The users can use weaker passwords as there is no custom password policy in place.
- No MFA enabled by default, the user has the choice to opt-out for the MFA authentication.

3.4.2 Honeypot with custom security policies but no MFA enabled

For the set-up of the second Azure AD honeypot, the attacks and interactions with the first honeypot are analyzed and the results are used to select the security measures that need to be in place to mitigate the number of successful attacks and increase the overall security posture.

Some of the security measures that are put in place to counteract the shortcomings in the first set-up are as below:

- The users are not able to invite the guest users, only the admin with guest invite role can do so. This stops the users with non-admin privileges from adding guest users to the directory.
- The users are no longer to see the details of other users in the same directory in the portal. This prevents the attacker from getting the complete user list from the portal.
  However, it important to note that the user still can get the list using Azure Powershell commands and no way can be prevented.

- A custom password policy is put in place which specifies that the minimum length of the password is 10 characters. Also, the password should contain at least one special character or the upper case alphabet in it

  To prevent the usage of the weaker and predictable passwords a dictionary of passwords is blocked and the users can not use any of those passwords, even though they meet the specified password policy criterion.

- The MFA is not enabled by default and the users can choose not to use the MFA authentication.

3.4.3 Honeypot with MFA enabled

During this set-up, very little modification is made to the second honeypot set-up. Multi-factor authentication is made mandatory by default so that every user has to opt for the MFA. Rest all the security measures put in place during the second honeypot set-up are kept the same.

We used this set-up to see the effect of the MFA compared to the second honeypot where MFA is not enabled.

3.5 Credential Leaking

To lure the attackers towards the honeypot setup, systematic leakage of the credentials on various web forums and websites is done. The credentials have been leaked in such a way that the attacker doesn't get red-flagged and suspicious of it.

To overcome this issue, Barron and Nikiforakis (2017) [38] claimed to have credentials of many accounts and leaking a test sample of 1 or a few accounts to authenticate their claim. This story makes it less suspicious and makes the leaked credentials valuable.

The credentials of the accounts which have no application access or administrative rights are leaked. They are used as the entry point to the system. Different Credentials are used for 9 different websites/forums on which the credentials were leaked. The different credentials were used to map the behavior of the attacker to the website where the credentials were leaked. Except for Pastebin, the credentials on the rest of the forums/websites were leaked manually due to reasons like posts being taken down or the account used to leak the credentials being blocked by the moderators.

The 9 different websites/forums used for the credential leakage are as below:

- Google docs : Bartjan Wolthuis
- Pastebin : Maurits-jan Oosterhof
- Hack Forums : Huibert Victorie
- Github Gist : Henkie Slaghuis
- Facebook Ethical Hacking Group 1 (More than 100K followers): Joep Ververda
- Facebook Ethical Hacking Group 2 (More than 100K followers): Tomas Euvelgunne
- Reddit- Microsoft Azure : Reinier Wieferink
- Reddit-Microsoft : Cathelijn Schuerman
- Reddit- hacking: security in practice: Mathijn Koendering

3.6 Retrieval of Logs

Once the attacker interacts with the honeypot setup, logs are generated. These logs provide a detailed description of the interaction of the attacker with the system along with the details of the attacker.

The two logs that are used in this thesis for data collection are:

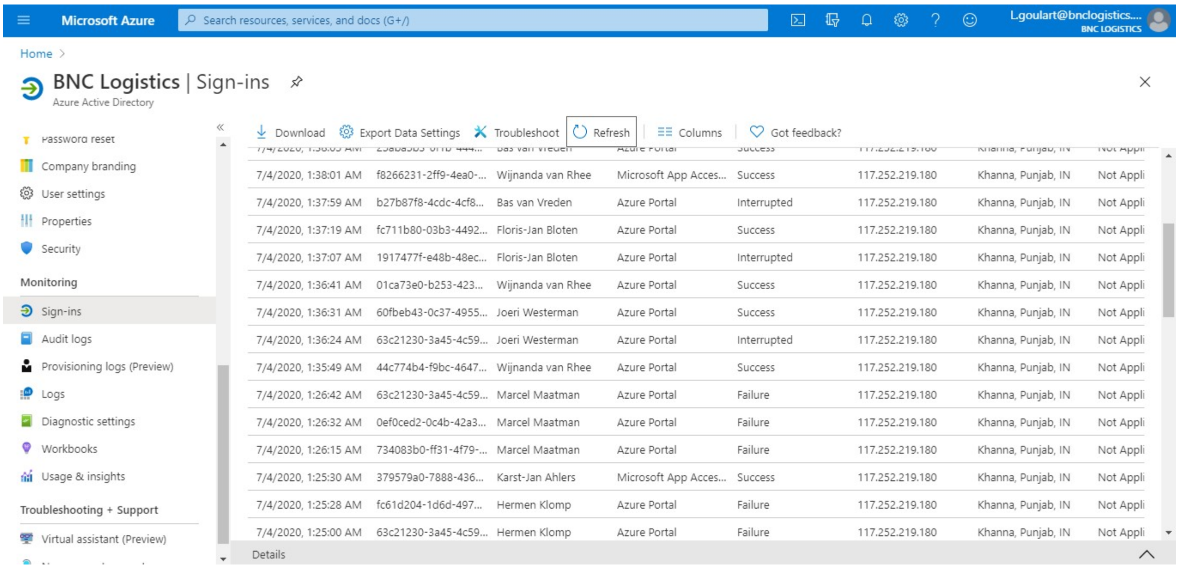
Sign- In Logs: These logs provide the user sign-in details and the information about the usage of the managed applications [39]. The default information in these logs are :

- Sign-in date
- User details
- Application accessed by user
- Sign-in status
- Risk detection status

- Multi-factor authentication status

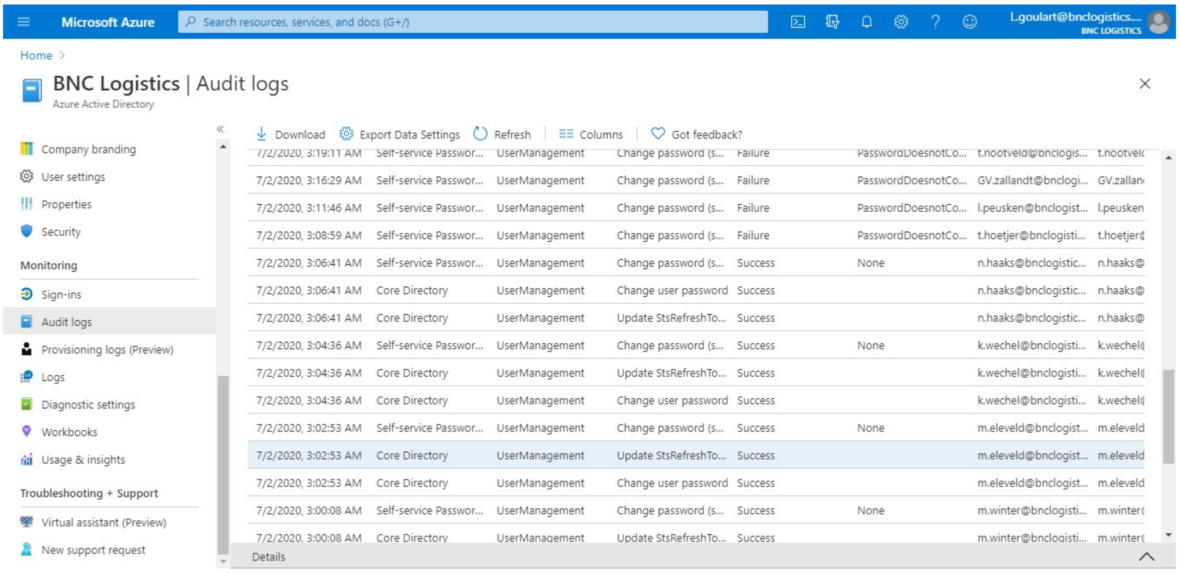
Audit Logs: These logs provide the information about the changes done to various features and services within the Azure AD [40]. The audit logs include changes made to any resource within the Azure AD like making changes to the users access policies, applications, etc.

- Date/time of occurrence
- Activity name/category
- Activity activity status/reason
- Activity target
- Activity initiator



7/4/2020, 1:38:05 AM	23aba5b5-0f1b-444...	Bas van Vreden	Azure Portal	Success	117.252.219.180	Khanna, Punjab, IN	Not Appli
7/4/2020, 1:38:01 AM	f8266231-2ff9-4ea0-...	Wijnanda van Rhee	Microsoft App Acces...	Success	117.252.219.180	Khanna, Punjab, IN	Not Appli
7/4/2020, 1:37:59 AM	b27b87f8-4cdc-4cf8...	Bas van Vreden	Azure Portal	Interrupted	117.252.219.180	Khanna, Punjab, IN	Not Appli
7/4/2020, 1:37:19 AM	fc711b80-03b3-4492...	Floris-Jan Bloten	Azure Portal	Success	117.252.219.180	Khanna, Punjab, IN	Not Appli
7/4/2020, 1:37:07 AM	1917477f-e48b-48ec...	Floris-Jan Bloten	Azure Portal	Interrupted	117.252.219.180	Khanna, Punjab, IN	Not Appli
7/4/2020, 1:36:41 AM	01ca73e0-b253-423...	Wijnanda van Rhee	Azure Portal	Success	117.252.219.180	Khanna, Punjab, IN	Not Appli
7/4/2020, 1:36:31 AM	60fbeb43-0c37-4955...	Joeri Westerman	Azure Portal	Success	117.252.219.180	Khanna, Punjab, IN	Not Appli
7/4/2020, 1:36:24 AM	63c21230-3a45-4c59...	Joeri Westerman	Azure Portal	Interrupted	117.252.219.180	Khanna, Punjab, IN	Not Appli
7/4/2020, 1:35:49 AM	44c774b4-f9bc-4647...	Wijnanda van Rhee	Azure Portal	Success	117.252.219.180	Khanna, Punjab, IN	Not Appli
7/4/2020, 1:26:42 AM	63c21230-3a45-4c59...	Marcel Maatman	Azure Portal	Failure	117.252.219.180	Khanna, Punjab, IN	Not Appli
7/4/2020, 1:26:32 AM	0ef0ced2-0c4b-42a3...	Marcel Maatman	Azure Portal	Failure	117.252.219.180	Khanna, Punjab, IN	Not Appli
7/4/2020, 1:26:15 AM	734083b0-ff31-4f79-...	Marcel Maatman	Azure Portal	Failure	117.252.219.180	Khanna, Punjab, IN	Not Appli
7/4/2020, 1:25:30 AM	379579a0-7888-436...	Karst-Jan Ahlers	Microsoft App Acces...	Success	117.252.219.180	Khanna, Punjab, IN	Not Appli
7/4/2020, 1:25:28 AM	fc61d204-1d6d-497...	Hermen Klomp	Azure Portal	Failure	117.252.219.180	Khanna, Punjab, IN	Not Appli
7/4/2020, 1:25:00 AM	63c21230-3a45-4c59...	Hermen Klomp	Azure Portal	Failure	117.252.219.180	Khanna, Punjab, IN	Not Appli

Figure 3.3: Sign-in logs generated



7/2/2020, 3:19:11 AM	Self-service Passwor...	UserManagement	Change password (s...	Failure	PasswordDoesnotCo...	t.nootveld@bnclogis...	t.nootvel
7/2/2020, 3:16:29 AM	Self-service Passwor...	UserManagement	Change password (s...	Failure	PasswordDoesnotCo...	GV.zallandt@bnclogi...	GV.zallan
7/2/2020, 3:11:46 AM	Self-service Passwor...	UserManagement	Change password (s...	Failure	PasswordDoesnotCo...	l.peusken@bnclogist...	l.peusken
7/2/2020, 3:08:59 AM	Self-service Passwor...	UserManagement	Change password (s...	Failure	PasswordDoesnotCo...	t.hoetjer@bnclogisti...	t.hoetjer@
7/2/2020, 3:06:41 AM	Self-service Passwor...	UserManagement	Change password (s...	Success	None	n.haaks@bnclogistic...	n.haaks@
7/2/2020, 3:06:41 AM	Core Directory	UserManagement	Change user password	Success		n.haaks@bnclogistic...	n.haaks@
7/2/2020, 3:06:41 AM	Core Directory	UserManagement	Update StsRefreshTo...	Success		n.haaks@bnclogistic...	n.haaks@
7/2/2020, 3:04:36 AM	Self-service Passwor...	UserManagement	Change password (s...	Success	None	k.wechel@bnclogisti...	k.wechel@
7/2/2020, 3:04:36 AM	Core Directory	UserManagement	Update StsRefreshTo...	Success		k.wechel@bnclogisti...	k.wechel@
7/2/2020, 3:04:36 AM	Core Directory	UserManagement	Change user password	Success		k.wechel@bnclogisti...	k.wechel@
7/2/2020, 3:02:53 AM	Self-service Passwor...	UserManagement	Change password (s...	Success	None	m.eleveld@bnclogist...	m.eleveld
7/2/2020, 3:02:53 AM	Core Directory	UserManagement	Update StsRefreshTo...	Success		m.eleveld@bnclogist...	m.eleveld
7/2/2020, 3:02:53 AM	Core Directory	UserManagement	Change user password	Success		m.eleveld@bnclogist...	m.eleveld
7/2/2020, 3:00:08 AM	Self-service Passwor...	UserManagement	Change password (s...	Success	None	m.winter@bnclogisti...	m.winter(
7/2/2020, 3:00:08 AM	Core Directory	UserManagement	Update StsRefreshTo...	Success		m.winter@bnclogisti...	m.winter(

Figure 3.4: Audit logs generated

These two logs help in answering the second research question of this thesis "what real-world attacks have been observed".These logs also provide information about the types of attacks done on the system and what real-world threats are observed. The logs can also provide information about the attacker and the tools he used while interacting with the honeypot.

After analyzing the collected data below information can be obtained :

- Whether the real-world attacks are present.
- Which security measure helps in counteracting against these attacks.
- Information about the Operating system, Geo-location, and tools used by the attacker

The sign-in and audit logs are used to visualize the interaction with the honeypot set-ups. Besides these logs, the Visitor's information collected from the website is used to have an overview of the path taken by the attacker. The audit logs and sign-in logs can be retrieved using Microsoft's reporting API. The Azure Active Directory (Azure AD) reporting API, allow users to have access to the data through RESTful API [41]. OAuth 2.0 protocol is used to authorize access to these APIs.

To retrieve the logs from the Azure active directory the user need to have one of the below admin roles:

- Security Reader

- Security Administrator
- Global Administrator

To retrieve the audit logs no premium license is required and the free tier accounts can use the API and collect the logs. However, for the sign-in logs, the premium license of P1 (or above) is required. In our case, we found that the sign-in logs generated are less than 250,000 and hence can be downloaded directly (manually) manually from the portal, without the need for the special premium license. Hence we retrieved only the audit logs generated using the reporting API and the sign-in logs are downloaded from the website (Azure AD portal). An example of sign-in and audit logs, showing their schema along with the different types of subscriptions for Azure AD is in the appendix.

3.7 Problems Encountered after deploying honeypot

During the deployment of the honeypot set, we faced some roadblocks that we had to get around so that the honeypot is accessible and visible to potential attackers. Below are the problems that we faced and the actions that we took to counter-act against them:

- **Password change by the user of entry accounts:** Microsoft Azure doesn't let the administrators the option of not letting the users change the password. So it is quite inevitable that the attacker won't change the password and stop other attackers from getting into the honeypot system.

  One way around this problem is to reset the password of all users after a fixed amount of time. This was achieved by using a PowerShell script that is being run on a virtual machine that resets the password after a set amount of time. The password is reset so that once the attacker changes the password during his interaction with the honeypot. The other user who might try to connect to the honeypot will also get access once it is reset.

  We selected the 600sec time so that it gives enough time for the attacker to interact without being prompted to type in the password again and also shorter so that the other attackers aren't getting blocked because the password has been changed by the user.

- **Problems during leaking the credentials:** During the leaking of the credentials several problems arose. The forums and the groups were removing the posts as they were containing private information like passwords and usernames. Posting these kinds of details and information is against most of the rules and guidelines set by the forums.

  The moderators and the admins of these groups were either removing the posts or in some cases were banning the accounts from posting anything new. In some portals like Pastebin, mentioning the word “username” or “password” would enable the captcha and hence the paste will only be made public once it was verified. This made it extremely impractical, as we planned to paste after every half an hour on Pastebin to create a consistent presence.

  To overcome these issues, was quite tedious and in some cases manual work. We need to create several accounts to post and words like “usrname”,”pwd” or “pass” were used. On some portals, an image was posted instead of the text as it was noticed that the pictures tend to get removed by the bot moderators less often than the text.

  The same thing was used for Pastebin, where “usrname” and “pwd” were used, and it didn’t activate the captcha. This made it possible to paste regularly and consistently without any intervention.

3.8 Attacks present during the three runs compared to the expected threats

Analyzing the attacks that are being done while interacting with the honeypot is not merely defined by the number of the sign-in logs or the number of unique IP addresses that accessed the system. We took a different approach and used the MITRE ATT&K matrices to identify the attacks that are conducted. The threats identified were already explained in section 3.1

During our preliminary analysis of the logs taken from the three different runs about different set-ups, we saw that almost all the predicted threats were present. The number of attacks although varied significantly based on the security measures put in place. This further justified our previous statement of the honeypot being attractive to the malicious users and they spent time and effort to interact and take advantage of the honeypot system.

In this chapter, we highlighted the different expected threats, various security postures of the organizations, and based on that we defined the three variations of the honeypot. Later sections of the chapter discussed the retrieval of the logs and the finding of the preliminary results from those logs. Additionally, we also mentioned the problems that were faced while setting up the honeypot system.

Chapter 4

Results per AAD Honeypot

This chapter deals with the analysis of the data obtained from the honeypots. The logs are parsed and information about the attacker and the type of attack conducted is obtained. This information both provides valid ground for the selection of security hardening measures as well as the data for the comparison of their effectiveness.

4.1 Results from Honeypot with default security measures

The credentials about the first honeypot set up were continuously leaked for 7 days from 29th July 2020 to 5th August 2020. During this period a total of 1588 individual sign-in log entries were generated from 47 unique IP addresses.
On 5th August 2020, the audit and sign-in logs were retrieved and the credentials leaking process was stopped. A Jupyter notebook was created to analyze and visualize the data obtained. The information obtained was used to see what kind of attacks were conducted and how the attacks vary based on geo-location.

The first and foremost thing to counter-check was whether the attacks mentioned in MITRE ATT&K matrices are present. We used these attacks as the basis for the attractiveness of the honeypot system.

In addition to this information collected from the logs help us in understanding how the active directory is attacked in the wild. This information can be quite helpful in designing future systems in such a way that they are resistant to them and hence have improved security and privacy.

4.1.1 Number of Password Changes

The attacker after successfully logging into the portal tries to change the password of the account to something else. The attacker tries to change the password usually so that he can continue to have access to the account. It was observed a total of 3 attackers out of 47 tried to change the password of the account.

The password change can be considered as a form of persistent attack as the attacker tries to have persistent access to the account, by changing the credentials of the account.

The IP addresses and the users that tried to change the password are as below:

IP address	User
175.XX.XX.156	Henkie Slaghuis
5.XX.XX.237	Mirre Eleveld
200.XX.XX.10	Mathijn Koendering

4.1.2 Number of Brute Force attempts

After analyzing the data we created a criterion that if there are ten consecutive failed sign-in attempts, we consider it as a brute force attack... In some cases, the attacker after attempting the password a certain number of times, waited for sometime ad tried again to brute force into the account. We considered them as separate brute force attempts.

Besides, we analyzed the applications the attacker uses for the brute force attack. The table below gives an overview of the accounts that the attacker tried to brute force, along with the IP address and the application used by the attacker.

IP address	User	Application	No. of Attempts
5.XX.XX.237	Lucas Goulart	Azure AD Powershell	1
85.XX.XX.105	Lucas Goulart	Azure AD Powershell	4
185.XX.XX.51	Lucas Goulart	Azure portal	1

It is interesting to see that only 1 user has been used for the brute force attack. This may be because this global admin has an obvious domain name ".onmicrosoft.com" in the principal username. The one with this username is usually the creator of the directory and has Global administrator privileges and other elevated accesses.

4.1.3 Password spray

This type of attack is a quite interesting and effective form of attack. This attack can be visualized by analyzing the failed sign-in attempts. Then the logs generated by that IP address are analyzed if it pertains to a single account or the username varies for each attempt. In case the username varies with every failed sign-in attempt, the attacker is trying to password spray instead of brute-forcing.

In our set up we found that 1 IP address tried to password spray 3 different times and he sprayed all the accounts in the directory.

IP address	Application	No. of attempts
5.XX.XX.237	Azure AD Powershell	3

4.1.4 Guest users

By default, the users in the Azure Active directory can invite the guest to the directory. They can do it for an individual guest user or send a bulk invite to multiple guest users. Attackers can utilize this feature to create guest accounts in the active directory and have a persistent presence in the directory.

In out set-up we found that 4 individual guest accounts were created and 1 bulk invite was sent by attacker. The guest users signed in a total of 10 times total . The attacker who created the accounts and the user he sent invite from are as below:

IP address	User	Guest User
85.XX.XX.105	Bartjan Wolthuis	aere455@gmail.com
200.XX.XX.10	Mathijn Koendering	beniazath@gmail.com
200.XX.XX.10	Mathijn Koendering	madrockstar99@gmail.com
63.XX.XX.52	Joep Ververda	peteraustin0031@gmail.com

Besides this one attacker sent a bulk guest invite but none of them registered and logged in . The details about the attacker who sent the bulk invite is as follows:

IP address	User	Bulk Invites Sent
200.XX.XX.10	Mathijn Koendering	1

4.1.5 Lateral Movement

Attackers usually do the lateral movement to get access to accounts that have elevated access or admin roles. To find the presence of lateral movement, the successful sign-in to the accounts whose credentials were not leaked is observed. A successful sign-in attempt means that the attacker was able to do the lateral movement i.e. the attacker was able to get the credentials of the account other than those which were leaked.

During the period for which the credentials were leaked, there was 1 successful lateral movement done. Different techniques used to get the credentials for lateral movement are brute force, password spray, social engineering, stealing login cookies, etc. However, in our case, the credentials were obtained by password spray.

The attacker was able to obtain the password of the below user. The details of the sign-in also reveal the application that the attacker tries to access after the lateral movement while interacting with the honeypot setup.

IP address	User	No of Total Sign Ins
5.XX.XX.237	Mirre Eleveld	SalesForce -3 Dropbox Business - 1 Azure AD powershell - 1 Azure Portal - 7

4.1.6 DOS attacks due to Lockouts

Smart lockouts are activated when the attacker enters the wrong credentials a certain specified number of times (by default the value is 10). Due to this, the accounts under lockout are inaccessible for a certain period to any user, whether it be the attacker or the intended user.

During the first run of honeypot set-up, a total of 129 sign-in attempts failed because of the lockouts. The accounts are under lockout because the attacker tried to brute force or password spray the accounts, causing them to be inaccessible for a certain amount of time. This causes a Denial of service to the actual users.

From the logs, we can find the IP address of the attacker which triggered the lockouts, and the accounts which were affected by that lockout.

IP address	User
85.XX.XX.105	Lucas Goulart
5.XX.XX.237	Jurrien Koetsie
5.XX.XX.237	Jan-Willem Braakhekke
5.XX.XX.237	Lucas Goulart
185.XX.XX.51	Lucas Goulart

4.1.7 Profiling of Attacker

In addition to the above information, we can get knowledge about the attacker. This knowledge can be useful in understanding the behavior of the attacker and creating access policies that can allow genuine users to log in while blocking the unwanted attackers from creating a security problem.

From the logs, we can get the details about the application which were used or the services that were accessed by the attacker, their location, information about their systems, the users they tried to sign in as and how many times they were successful, and how many times they failed.

Location of Attackers: From the sign-in logs we can get the information about the Geo-location of the attacker. This information is really important in securing the Azure Ad and resources from attackers, by blocking access from certain geo-locations. An access policy allowing access from certain locations and blocking access from other locations can help in securing the Azure AD if the operations of the company are localized and not widespread throughout the globe.

Country	No of Unique IPs
USA	14
Sri Lanka	4
Philippines	4
India	3
Netherlands	2
Argentina	2
UK	2
Egypt	2
Iceland	2
France	2

Country	No of Unique IPs
Romania	1
Morocco	1
Mexico	1
Greece	1
Pakistan	1
Tunasia	1
Australia	1
Italy	1
Brazil	1
Russia	1

Users Accessed: Below is the information about the number of times the attempts were made to sign in to that account. The information about the number of successful and failed attempts is also included. The information besides the created users also includes the guest users that were created by the attackers. From this information we can see which account is the most target account and is the prime target for the attackers.

User	Failed Signins	Sucess Signins	No. of Unique Ips
Bartjan Wolthuis	20	72	20
Maurits-jan Oosterhof	25	29	11
Huibert Victorie	28	21	11
Henkie Slaghuis	30	15	5
Joep Ververda	23	8	5
Tomas Euvelgunne	22	39	4
Lucas Goulart	139	0	3
Wijnanda van Rhee	32	0	3
Mathijn Koendering	25	11	2
Cathelijn Schuerman	26	9	2
Reinier Wieferink	23	2	2
aere455g	0	4	1
Scare Crow	0	4	1
beniazath	0	2	1
Mirre Eleveld	27	12	1
Bas van Vreden	26	0	1

User	Failed Signins	Sucess Signins	No. of Unique Ips
Fransien Netters	26	0	1
Evelien Hoekjan	26	0	1
Emiel Vlogtman	25	0	1
Gust van ’t Zallandt	25	0	1
Floris-Jan Bloten	25	0	1
Hermen Klomp	25	0	1
Eefje van Breen	25	0	1
Niek Haaks	24	0	1
Harm Klein Jans	24	0	1
Kevin Langkamp	24	0	1
Joren Geertsen	24	0	1
Mannes Kuilaard	24	0	1
Linneke Hazelaar	24	0	1
Marcel Meinders	24	0	1
Jeroen Podt	24	0	1
Jan-Willem Braakhekke	24	0	1
Lennert Koenen	24	0	1
Marcel Reefhuis	24	0	1
Joeri Westerman	24	0	1
Margriet Nijenhuis	24	0	1
Kees te Wechel	24	0	1
Luuk Peusken	23	0	1
Roelof Wielents	23	0	1
Ron Slomp	23	0	1
Renske Pool	23	0	1
Mirre Winter	23	0	1
Marcel Maatman	23	0	1
Quintijn Weinreich	23	0	1
Jurrijn ten Brinke Degener	23	0	1
Karst-Jan Ahlers	23	0	1
Michiel Tijselink	23	0	1
Jurrien Koetsie	22	0	1
Sofietje van Mulder	22	0	1
Louw Apperlo	22	0	1
Sijbrand Boddeman	22	0	1
Rosemarije Heupers	21	0	1

User	Failed Signins	Sucess Signins	No. of Unique Ips
Toon Hoetjer	21	0	1
Sijbrand Fijneman	21	0	1
Trijntje Nootveld	20	0	1

Operating system and Browser: The information about the operating system and the browser (if the attacker didn't use Azure Active directory powershell) the attacker used to sign in and interact with the honeypot is also in the sign in logs.
The table below shows the different operating systems and browsers that were used.

Operating System	No. of Unique Ips
Windows	28
Android	16
iOS	3

Browser	No. of Unique Ips
Chrome Windows	26
Chrome Mobile	14
Edge / IE	2
Safari	3

Applications or Resources accessed: The logs provided important information on whether the applications registered with the azure active directory are being accessed or tried to be accessed. During the first set-up, the 2 applications out of 3 were being accessed. This also provides information about the resource the attacker interacted with. Below is the table having the resource/application name along with the number of IPs that interacted with it

Application Name	No. of Unique Ips
Azure Web Portal	43
Azure Active Directory PowerShell	2
Dropbox	2
Salesforce	2

4.2 Results from Honeypot with custom security policies but no MFA enabled

After the analysis of the first honeypot set-up, the security measures to counteract against the attacks and threats were included in the deployed honeypot. The Credentials after the security hardening measures were leaked on the same portals and forums. The credentials were leaked in the same manner as in the first set-up for 7 days, from 24th August 2020 to 31st Sept 2020. During this period a total of 42 unique IP addresses generated a total of 771 sign-ins.

The Jupyter notebook was again utilized to analyze the data compare it with the first set-up. Based on this comparison the effectiveness of the security hardening measures was gauged.

4.2.1 Number of Password Changes

As there is no way by which the administrator can set up the users from changing the password of their accounts. The users were able to change the password of their accounts in the second honeypot set-up as well. During the second set-up, a total of 6 attackers changed the password of the account. in which they were logged in. The IP addresses and the users that tried to change the password are as below

IP address	User
94.XX.XX.56	Maurits-jan Oosterhof
195.XX.XX.8	Bartjan Wolthuis
200.XX.XX.93	Cathelijn Schuerman
2.XX.XX.13	Maurits-jan Oosterhof
103.XX.XX.204	Bartjan Wolthuis
41.XX.XX.164	Tomas Euvelgunne

4.2.2 Number of Brute Force attempts

As the access to the Azure portal was restricted to the users. The users couldn't see the users in the active directory. The only way the attacker could see the list of the users and their roles is by using the Azure Powershell. Due to this, no attacker met the criterion of 10 consecutive failed sign-ins for the same user, and hence there has been no attempt of brute-forcing into another account during the second set-up.

4.2.3 Password Spray

During the second deployment of the honeypot set-up, 1 attacker tried to password spray multiple accounts to gain access to the privileged accounts. No successful password was guessed during the password spray.
Below is the number no of times the attacker tried to do the password spray and the application used for the password spray.

IP address	Application	No. of attempts
194.XX.XX.20	Azure AD Powershell	1

4.2.4 Guest users

Due to the users being stripped of the privileges to invite the external guest users. During the second set-up, no guest users were invited. Also due to the above security policy, no bulk guest invites were sent.

4.2.5 Lateral Movement

The use of custom password policies forces users to use the password of higher difficulty levels and avoid usage of weaker predictable passwords. This helps in preventing the attacker from laterally moving from one account to another.
During the second honeypot set-up run, no lateral movements were observed. One user tried to spray attack accounts multiple times, however, due to stronger passwords being used, it was quite difficult to predict the correct password.

4.2.6 DOS attacks due to Lockouts

Due to stricter access policies, the attacker was able to get far less information from the Azure web portal compared to the first deployment run. This made it less obvious to the attacker to have information related to the active directory. Information related to users, admins, enterprise application registered, users having access to that application is not accessible on the portal.

These measures resulted in a lesser number of password sprays or brute force attacks conducted on the active directory. This resulted in a lesser number of times the accounts were locked out.

We observed that during the second run only 1 time the account was inaccessible due to lockouts. The details regarding the attacker and his IP address along with the user who was locked-out is given below in the table:

IP address	User
194.XX.XX.20	Lucas Goulart

4.2.7 Profiling of Attacker

The details about the attackers who interacted with the honeypot during the 2nd deployment run are given below:

Location of Attackers: The Location of the attackers who interacted with the honeypot set-up during the 2nd run is obtained from the sign-in logs.

The table below shows the Geo-location of the attackers along with number of unique IPs from that country.

Country	No of Unique IPs
USA	10
UK	7
Australia	3
Russia	3
Iceland	2
Denmark	2
Turkey	1
Portugal	1
Serbia	1
Germany	1
Argentina	1

Country	No of Unique IPs
India	1
Austria	1
Mexico	1
Sweden	1
Brazil	1
France	1
Chile	1
Morocco	1
Canada	1
Netherlands	1

Users Accessed: The user accounts that the attacker user to interact with the set-up or tried to login along with both successful and failed attempts are shown in the table below.

User	Failed Signins	Sucess Signins	No. of Unique Ips
Bartjan Wolthuis	24	68	18
Maurits-jan Oosterhof	18	41	9
Huibert Victorie	4	29	8
Tomas Euvelgunne	11	17	2
Reinier Wieferink	11	8	3
Cathelijn Schuerman	12	7	2
Joep Ververda	12	7	3
Mathijn Koendering	11	5	2
Henkie Slaghuis	11	4	3
Eefje van Breen	12	0	1
Emiel Vlogtman	12	0	1
Lucas Goulart	12	0	1
Bas van Vreden	12	0	1
Sijbrand Boddeman	11	0	1
Michiel Tijselink	11	0	1
Marcel Meinders	11	0	1
Roelof Wielents	11	0	1
Gust van ’t Zallandt	11	0	1
Luuk Peusken	11	0	1
Jeroen Podt	11	0	1
Floris-Jan Bloten	11	0	1
Jurrijn ten Brinke Degener	11	0	1
Renske Pool	11	0	1
Lennert Koenen	11	0	1
Jurrien Koetsie	11	0	1
Rosemarije Heupers	11	0	1
Fransien Netters	11	0	1
Marcel Maatman	11	0	1
Mirre Eleveld	11	0	1
Ron Slomp	11	0	1
Kees te Wechel	11	0	1
Evelien Hoekjan	11	0	1
Jan-Willem Braakhekke	11	0	1
Louw Apperlo	11	0	1
Marcel Reefhuis	11	0	1
Joeri Westerman	11	0	1
Hermen Klomp	11	0	1

User	Failed Signins	Sucess Signins	No. of Unique Ips
Niek Haaks	11	0	1
Linneke Hazelaar	11	0	1
Harm Klein Jans	11	0	1
Quintijn Weinreich	11	0	1
Karst-Jan Ahlers	11	0	1
Margriet Nijenhuis	11	0	1
Mannes Kuilaard	11	0	1
Sijbrand Fijneman	11	0	1
Mirre Winter	11	0	1
Joren Geertsen	11	0	1
Kevin Langkamp	11	0	1
Sofietje van Mulder	10	0	1
Toon Hoetjer	10	0	1
Wijnanda van Rhee	9	0	1
Trijntje Nootveld	9	0	1

Operating system and Browser: The operating system along with the browser information of the attackers during the second phase are in the table:

Operating System	No. of Unique Ips
Windows	36
Android	4
iOS	2

Browser	No. of Unique Ips
Chrome Windows	28
Chrome Mobile	4
Edge / IE	2
Safari	2

Applications or Resources accessed: The application or resources registered with the azure active directory, that the attacker tried or accessed is shown in the table below:

Application Name	No. of Unique Ips
Azure Web Portal	42
Azure Active Directory PowerShell	1

4.3 Results from Honeypot with MFA enabled

The final Setup was deployed with the Multi-Factor Authentication enabled. It was deployed to see how effective the MFA is in hardening the security posture of the honeypot setup. It was deployed for 7 days between 7th September 2020 to 14th Sept 2020. The credentials were leaked in the same manner as during the first two setups.

During this period a total of 37 unique IPs tried to access the honeypot setup, but all of them failed because of the MFA being enabled on them.

4.3.1 Number of Password Changes

The attackers can change the password of the user accounts in which they are signed in. However during the third setup, due to MFA, all the sign-in attempts failed. Due to this, the attackers were not able to change the passwords of the users in the honeypot setup. Hence during the period for which the honeypot was deployed no password changes were made.

4.3.2 Number of Brute Force attempts Password Spray

As the attackers were not able to sign in, getting the information about the users in the active directory is impossible. Hence due to this no Brute force or password spray attempts were made on other users during the period for which honeypot was deployed.

4.3.3 Guest users

The attackers could invite the guest users only if they can properly sign in to the directory. However due to the MFA being enabled the attackers weren't able to get into the active directory, hence no guest users were invited during the period of honeypot deployment.

4.3.4 Lateral Movement

The MFA didn't allow the attacker to sign in even if he has the login credentials. This prevented him from getting into the active directory honeypot and stopped the attacker from laterally moving from one account to another completely. Hence we

observed no lateral movements.
During the second honeypot set-up run, no lateral movements were observed. One user tried to spray attack accounts multiple times, however, due to stronger passwords being used, it was quite difficult to predict the correct password.

4.3.5 DOS attacks due to Lockouts

During the third honeypot setup, due to the absence of password spray or brute force attack, the attackers were not able to lock out any user account.

During the deployment period, no lockout of any account was observed and all user accounts were accessible at all times.

4.3.6 Profiling of Attacker

The details about the attackers who interacted with the honeypot during the 3rd deployment run are given below:

Location of Attackers: The Location of the attackers who interacted with the honeypot set-up during the 3rd run is obtained from the sign-in logs.

The table below shows the Geo-location of the attackers along with several unique IPs from that country.

Country	No of Unique IPs
USA	8
Russia	5
Australia	3
France	3
India	3
Germany	2
Netherlands	2
Iceland	2
UK	2
Greece	1
Brazil	1
Denmark	1
Turkey	1
Serbia	1
Mexico	1
Portugal	1

Users Accessed: The user accounts that the attacker user to interact with the set-up or tried to login along with both successful and failed attempts are shown in the table below.

User	Failed Signins	Sucess Signins	No. of Unique Ips
Bartjan Wolthuis	9	0	9
Maurits-jan Oosterhof	2	0	2
Huibert Victorie	5	0	5
Tomas Euvelgunne	4	0	4
Reinier Wieferink	3	0	3
Cathelijn Schuerman	5	0	5
Joep Ververda	3	0	3
Mathijn Koendering	5	0	5
Henkie Slaghuis	1	0	1
Rest of users	0	0	0

Operating system and Browser: The operating system along with the browser information of the attackers during the second phase are in the table:

Operating System	No. of Unique Ips
Windows	28
Android	6
iOS	3

Browser	No. of Unique Ips
Chrome Windows	23
Chrome Mobile	6
Edge / IE	5
Safari	3

Applications or Resources accessed: The application or resources registered with the azure active directory, that the attacker tried or accessed is shown in the table below:

Application Name	No. of Unique Ips
Azure Web Portal	35
Azure Active Directory PowerShell	2

4.4 Comparison between the three honeypot setups

Based on the information that we got from the data collected during the 3 honeypot setups, we can analyze the hardening measures put in place to improve the security posture of the Azure Active Directory. In addition to that, we were also able to see what tools the attacker uses to interact with the honeypot setup and also information about the attacker itself.

The three different setups showed how much difference the hardening measures makeover the default setting of the active directory. We have observed that using default and better security measures drastically reduces the threat vectors and secure the overall system considerably.

The table below gives an overview of how the different threats identified in section 3.2 can be mitigated or reduced by these hardening measures.

Attack / Threats	Number of incidences		
	First Setup	Second set up	Third Setup
Password Change	3	6	0
Brute Force	6	0	0
Password Spray	3	1	0
Lateral Movement	1	0	0
DOS Attack / Lockout	5	1	0
Guest User Creation	4	0	0

The table clearly shows the superiority of the MFA and how much efficient it is in stopping the attacks against the active directory. It completely blocked all the identified threats and secured the setup completely.

Comparing the first 2 setups we can see that having custom security policies do provide some positive impact on the security of the system. We can see from the table above that the threats like account creation and lateral movements have been completely stopped. Threats like Brute force/ password spray and DOS attack has also been reduced, this is because the attacker is not able to visualize the information about the directory and other users easily compared to the first honeypot setup.

4.5 Attacks that can't be prevented.

The attack threats like account manipulation remain the same, as in our case the attacker is still able to change the password of the logged-in account and hence can result in the persistent presence of the attacker. This is because the administrator can not stop users who are logged in from changing the password of the accounts. The Azure AD by default has some really serious security shortcomings which can pave an easy entry point for the attacker to the system.

- Attacker can get the list of the users and other information about the directory, even if it is made hidden using Azure AD Powershell
- Administrators cant disallow users from changing the password of their accounts
- Weaker passwords are being accepted via the Powershell even though the Azure Portal rejects the same. Allowing users to use weaker predictable passwords.

This chapter provided a detailed overview of the attacks that were encountered during the three different runs. The number of attacks was used to evaluate the performance of the security measures and their efficacy against each other. We were also able to find certain threats which loom irrespective of the security measure chosen (in case MFA is disabled).

Chapter 5

Conclusions and Future Work

The goal of the thesis was to extend the use of the honeypot to the Azure Active Directory (Azure AD) and to evaluate the performance of security hardening measures. To achieve the goals of the thesis we intend to answer the three research goals defined. These help in planning to deploy and evaluating the honeypots setups with different security measures. Below are the questions that are the point of interest in this thesis:

- What criteria define an attractive Azure AD honeypot set-up and how they are met.
- What different honeypot set-ups are deployed and what real-world attacks have been observed.
- How the various security hardening measures perform against these attacks observed.

The first research question was answered by identifying the criteria that define an attractive honeypot system (section 2.2). These criteria provide a guideline for establishing a framework for the proposed honeypot system. To meet these criteria we created a fake website, used Dutch names for the employees, and tried to complete the profile details of the employees of the fake company. We linked enterprise applications like Workplace, Dropbox, and Salesforce to the Azure AD, as they are quite relevant to the logistics company and give a false sense of reward to the attacker. To ensure the ethical and legal obligations of the honeypots are met, we reset the set-up after every 600sec, restricted access of all the users, and also anonymized the personal information of all the attackers. The proposed honeypot system was able to meet all those requirements and was deemed attractive.

To answer the second research question, we first identified the three broadly defined security postures. These are based on how the organizations configure their

Azure AD in the real-world based on various factors. Using this information we deployed three variations of the honeypot systems and intend to identify the real-world threats that are observed. The three set-ups chosen are with the default settings, with a custom security policy but no MFA enabled, and with MFA enabled. The credentials about the set-ups were leaked on various websites, forums, and Facebook groups related to hacking, cloud infrastructure, and Azure AD. The credentials were leaked over 7 days for each set-up.

To analyze the results audit and sign-in logs were used. Sign-in and audit logs retrieved from the various versions of the honeypot system deployed reveal that the system was attractive and the attackers were not able to identify it as a fake system/honeypot. They were interacting with the system and were trying to gain access to resources that were not accessible earlier to them. We were also able to find out the presence of the real-world threats as defined by the MITRE ATT&CK matrix. These threats are exactly what the real world systems experience and the presence of those means that for the attacker, honeypot system is indistinguishable from the real system.

Finally to answer the third research question the logs from the various honeypots were collected and parsed for data. During the analysis of the data, we were able to find out how the various security hardening measures compare to each other and the default settings of the Azure AD (section 4.4). The logs from these setups reveal that by default several security measures need to be put in place to have a securer Azure AD for identity and access management. We also found out that the MFA is the most securer and efficient way to harden the security of Azure AD. This is inline what the industry expectations and the real-world scenarios as well. However, in absence of MFA certain security measures do benefit and make the Azure AD more resilient to many threat vectors like brute force, lateral movement, account creation and password spray.

Additionally from our research, we were able to identify certain shortcomings in the security of the Azure AD which can lead to an attack. These shortcomings can be easily exploited by the attacker to gain easy access to the Azure AD. We have identified three such issues in section 4.5 of this thesis.

During the research, we observed that the number of attackers (identified by unique IPs) is lower, but that is because the information about the leaked credentials is regularly being taken down by the administrators and moderators of the forums/websites. In some cases leaking this confidential information leads to the banning of the account used to leak the credentials. This reduced the overall presence of the credentials on the forums/websites and very few people very able to see it, and hence the lower number. However, it can be argued that instead of the lower numbers the results are quite in line with what one should expect and that makes absolute sense.

All the logs collected and the script used to reset the honeypot are stored in the private Github repository for privacy reasons. In case anyone wants to have access to the data and scripts, they can do that after gaining access to the repository. `https://github.com/AtifMKhan/AD`

The work done in this thesis pave the way for using the honeypots in IAM solutions like Azure AD. In future researches the problem of lower interactions can be solved, if the credentials are leaked in a way that they are accessible and available on the internet for a prolonged duration of time.

Bibliography

[1] Azure active directory. [Online]. Available: https://azure.microsoft.com/en-in/services/active-directory/

[2] 99% of compromised microsoft enterprise accounts lack multi-factor authentication. [Online]. Available: https://nakedsecurity.sophos.com/2020/03/09/99-of-compromised-microsoft-enterprise-accounts-lack-mfa/3

[3] 44 million azure ad/microsoft accounts compromised; password problems highlighted. [Online]. Available: https://www.scmagazineuk.com/44-million-azure-ad-microsoft-accounts-compromised-password-problems-highlighted/article/166813

[4] A password spraying tool for microsoft online account. [Online]. Available: https://github.com/dafthack/MSOLSpray

[5] L. Spitzner, *Honeypots: Tracking Hackers*. USA: Addison-Wesley Longman Publishing Co., Inc., 2002.

[6] L. Spitzner, "The honeynet project: trapping the hackers," *IEEE Security Privacy*, vol. 1, no. 2, pp. 15–23, 2003.

[7] Tom liston talks about labrea. [Online]. Available: https://labrea.sourceforge.io/Intro-History.html

[8] N. Krawetz, "Anti-honeypot technology," *IEEE Security and Privacy*, vol. 2, no. 1, p. 76–79, Jan. 2004. [Online]. Available: https://doi.org/10.1109/MSECP.2004.1264861

[9] Changing the game: The art of deceiving sophisticated attackers. [Online]. Available: https://www.ccdcoe.org/uploads/2018/10/d2r2s6_serrano.pdf

[10] Honeypots in network security. [Online]. Available: http://www.diva-portal.org/smash/get/diva2:327476/fulltext01.

[11] R. M. D. W. I. A. by How They Are Hacking:A Classification of Current Remote Desktop Modus Operandi, "https://davidebove.com/files/thesis-bove-public.pdf," in *)*, 2018.

[12] S. Brown, R. Lam, S. Prasad, S. Ramasubramanian, and J. Slauson, "Honeypots in the cloud," 2012.

[13] D. Bove and T. Müller, "Investigating characteristics of attacks on public cloud systems," in *2019 6th IEEE International Conference on Cyber Security and Cloud Computing (CSCloud)/ 2019 5th IEEE International Conference on Edge Computing and Scalable Cloud (EdgeCom)*, June 2019, pp. 89–94.

[14] N. Provos, "Honeyd: A virtual honeypot daemon (extended abstract)," 01 2003.

[15] What's the difference between a high interaction honeypot and a low interaction honeypot? [Online]. Available: https://www.guardicore.com/2019/1/high-interaction-honeypot-versus-low-interaction-honeypot/

[16] N. Antonopoulos and L. Gillam, *Cloud computing*. Springer, 2010.

[17] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "A view of cloud computing," *Commun. ACM*, vol. 53, no. 4, p. 50–58, Apr. 2010. [Online]. Available: https://doi.org/10.1145/1721654.1721672

[18] The nist definition of cloud computing. [Online]. Available: https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-145.pdf

[19] Single sign-on. [Online]. Available: https://en.wikipedia.org/wiki/Single_sign-on

[20] Gdpr personal data. [Online]. Available: https://gdpr-info.eu/issues/personal-data/

[21] M. J. . H. M. Sokol, P., "Honeypots and honeynets: issues of privacy," *EURASIP Journal on Information Security*, Feb. 2017. [Online]. Available: https://doi.org/10.1186/s13635-017-0057-4

[22] R. N. Dahbul, C. Lim, and J. Purnama, "Enhancing honeypot deception capability through network service fingerprinting," *Journal of Physics: Conference Series*, vol. 801, p. 012057, jan 2017. [Online]. Available: https://doi.org/10.1088/1742-6596/801/1/012057

[23] What is a honeypot? [Online]. Available: https://www.kaspersky.co.in/resource-center/threats/what-is-a-honeypot

[24] Honeypots revealed? [Online]. Available: http://www.infosecwriters.com/text_resources/pdf/Honeypots.pdf

[25] What is application management? [Online]. Available: https://docs.microsoft.com/en-us/azure/active-directory/manage-apps/what-is-application-management

[26] Salesforce. [Online]. Available: https://www.salesforce.com/

[27] Workplace. [Online]. Available: https://www.workplace.com/

[28] Dropbox business. [Online]. Available: https://www.dropbox.com/business

[29] T. A. L. John D. Howard, "A common language for computer security incidents," 1998. [Online]. Available: https://prod.sandia.gov/techlib-noauth/access-control.cgi/1998/988667.pdf

[30] Azure ad matrix. [Online]. Available: https://attack.mitre.org/matrices/enterprise/cloud/azuread/

[31] Credentials processes in windows authentication. [Online]. Available: https://docs.microsoft.com/en-us/windows-server/security/windows-authentication/credentials-processes-in-windows-authentication

[32] Study reveals average person has 100 passwords. [Online]. Available: https://tech.co/news/average-person-100-passwords

[33] E. Stobert and R. Biddle, "The password life cycle: User behaviour in managing passwords," in *10th Symposium On Usable Privacy and Security (SOUPS 2014)*. Menlo Park, CA: USENIX Association, Jul. 2014, pp. 243–255. [Online]. Available: https://www.usenix.org/conference/soups2014/proceedings/presentation/stobert

[34] Nist special publication 800-63b. [Online]. Available: https://pages.nist.gov/800-63-3/sp800-63b.html

[35] Multi-factor authentication. [Online]. Available: https://en.wikipedia.org/wiki/Multi-factor_authentication

[36] Survey of 2600 it pros: Password procedures still are a cyber security fail. [Online]. Available: https://blog.knowbe4.com/survey-of-2600-it-pros-password-procedures-still-are-a-cyber-security-fail

[37] How to overcome the security hate factor when implementing mfa. [Online]. Available: https://blog.identityautomation.com/how-to-overcome-the-security-hate-factor-when-implementing-mfa

[38] T. Barron and N. Nikiforakis, “Picky attackers: Quantifying the role of system properties on intruder behavior,” in *Proceedings of the 33rd Annual Computer Security Applications Conference*, ser. ACSAC 2017. New York, NY, USA: Association for Computing Machinery, 2017, p. 387–398. [Online]. Available: https://doi.org/10.1145/3134600.3134614

[39] Sign-in activity reports in the azure active directory portal. [Online]. Available: https://docs.microsoft.com/en-us/azure/active-directory/reports-monitoring/concept-sign-ins

[40] Audit activity reports in the azure active directory portal. [Online]. Available: https://docs.microsoft.com/en-us/azure/active-directory/reports-monitoring/concept-audit-logs

[41] Get started with the azure active directory reporting api. [Online]. Available: https://docs.microsoft.com/en-us/azure/active-directory/reports-monitoring/concept-reporting-api

Appendix A

Appendix

A.1 Example of Audit logs

The audit logs retrieved from the Azure Directory using the reporting API is in JSON format and has below schema and contains the below mentioned information:

```
{
    "time": "2019-03-12T16:02:15.5522137Z",
    "resourceId": "/tenants/<TENANT ID>/providers/Microsoft.aadiam",
    "operationName": "Sign-in activity",
    "operationVersion": "1.0",
    "category": "SignInLogs",
    "tenantId": "<TENANT ID>",
    "resultType": "50140",
    "resultSignature": "None",
    "resultDescription": "This error occurred due to 'Keep me signed in'
    interrupt when the user was signing-in.",
    "durationMs": 0,
    "callerIpAddress": "<CALLER IP ADDRESS>",
    "correlationId": "a75a10bd-c126-486b-9742-c03110d36262",
    "identity": "Timothy Perkins",
    "Level": 4,
    "location": "US",
    "properties":
        {
            "id":"0231f922-93fa-4005-bb11-b344eca03c01",
            "createdDateTime":"2019-03-12T16:02:15.5522137+00:00",
            "userDisplayName":"Timothy Perkins",
            "userPrincipalName":"<USER PRINCIPAL NAME>",
            "userId":"<USER ID>",
            "appId":"<APPLICATION ID>",
            "appDisplayName":"Azure Portal",
            "ipAddress":"<IP ADDRESS>",
            "status":
            {
                "errorCode":50140,
                "failureReason":"This error occurred due to 'Keep me signed in'
                interrupt when the user was signing-in."
            },
            "clientAppUsed":"Browser",
            "deviceDetail":
            {
                "operatingSystem":"Windows 10",
                "browser":"Chrome 72.0.3626"
            },
            "location":
                {
                    "city":"Bellevue",
                    "state":"Washington",
                    "countryOrRegion":"US",
                    "geoCoordinates":
                    {
                        "latitude":45,
                        "longitude":122
                    }
                },
            "correlationId":"a75a10bd-c126-486b-9742-c03110d36262",
            "conditionalAccessStatus":"notApplied",
            "appliedConditionalAccessPolicies":
            [
                {
                    "id":"ae11ffaa-9879-44e0-972c-7538fd5c4d1a",
                    "displayName":"Hr app access policy",
                    "enforcedGrantControls":
                    [
                        "Mfa"
                    ],
                    "enforcedSessionControls":
                    [
                    ],
                    "result":"notApplied"
                },
                {
                    "id":"b915a70b-2eee-47b6-85b6-ff4f4a66256d",
                    "displayName":"MFA for all but global support access",
                    "enforcedGrantControls":[],
                    "enforcedSessionControls":[],
                    "result":"notEnabled"
                },
                {
                    "id":"830f27fa-67a8-461f-8791-635b7225caf1",
                    "displayName":"Header Based Application Control",
                    "enforcedGrantControls":["Mfa"],
                    "enforcedSessionControls":[],
                    "result":"notApplied"
                },
                {
                    "id":"8ed8d7f7-0a2e-437b-b512-9e47bed562e6",
                    "displayName":"MFA for everyones",
                    "enforcedGrantControls":[],
                    "enforcedSessionControls":[],
                    "result":"notEnabled"
                },
                {
                    "id":"52924e0f-798b-4afd-8c42-49055c7d6395",
                    "displayName":"Device compliant",
                    "enforcedGrantControls":[],
                    "enforcedSessionControls":[],
                    "result":"notEnabled"
                },
             ],
            "isInteractive":true,
            "tokenIssuerType":"AzureAD",
            "authenticationProcessingDetails":[],
            "networkLocationDetails":[],
            "processingTimeInMilliseconds":0,
            "riskDetail":"hidden",
            "riskLevelAggregated":"hidden",
            "riskLevelDuringSignIn":"hidden",
            "riskState":"none",
            "riskEventTypes":[],
            "resourceDisplayName":"windows azure service management api",
            "resourceId":"797f4846-ba00-4fd7-ba43-dac1f8f63013",
            "authenticationMethodsUsed":[]
        }
}
```

A.2 Example of Sign in logs:

The sign in logs generated are downloaded in csv or JSON format from the website and has below schema:

```
{
    "time": "2019-03-12T16:02:15.5522137Z",
    "resourceId": "/tenants/<TENANT ID>/providers/Microsoft.aadiam",
    "operationName": "Sign-in activity",
    "operationVersion": "1.0",
    "category": "SignInLogs",
    "tenantId": "<TENANT ID>",
    "resultType": "50140",
    "resultSignature": "None",
    "resultDescription": "This error occurred due to 'Keep me signed in' interrupt when the user was
    "durationMs": 0,
    "callerIpAddress": "<CALLER IP ADDRESS>",
    "correlationId": "a75a10bd-c126-486b-9742-c03110d36262",
    "identity": "Timothy Perkins",
    "Level": 4,
    "location": "US",
    "properties":
        {
            "id":"0231f922-93fa-4005-bb11-b344eca03c01",
            "createdDateTime":"2019-03-12T16:02:15.5522137+00:00",
            "userDisplayName":"Timothy Perkins",
            "userPrincipalName":"<USER PRINCIPAL NAME>",
            "userId":"<USER ID>",
            "appId":"<APPLICATION ID>",
            "appDisplayName":"Azure Portal",
            "ipAddress":"<IP ADDRESS>",
            "status":
            {
                "errorCode":50140,
                "failureReason":"This error occurred due to 'Keep me signed in'
                interrupt when the user was signing-in."
            },
            "clientAppUsed":"Browser",
            "deviceDetail":
            {
                "operatingSystem":"Windows 10",
                "browser":"Chrome 72.0.3626"
            },
            "location":
                {
                    "city":"Bellevue",
                    "state":"Washington",
                    "countryOrRegion":"US",
                    "geoCoordinates":
                    {
                        "latitude":45,
                        "longitude":122
                    }
                },
            "correlationId":"a75a10bd-c126-486b-9742-c03110d36262",
            "conditionalAccessStatus":"notApplied",
            "appliedConditionalAccessPolicies":
            [
                {
                    "id":"ae11ffaa-9879-44e0-972c-7538fd5c4d1a",
                    "displayName":"Hr app access policy",
                    "enforcedGrantControls":
                    [
                        "Mfa"
                    ],
                    "enforcedSessionControls":
                    [
                    ],
                    "result":"notApplied"
                },
                {
                    "id":"b915a70b-2eee-47b6-85b6-ff4f4a66256d",
                    "displayName":"MFA for all but global support access",
                    "enforcedGrantControls":[],
                    "enforcedSessionControls":[],
                    "result":"notEnabled"
                },
                {
                    "id":"830f27fa-67a8-461f-8791-635b7225caf1",
                    "displayName":"Header Based Application Control",
                    "enforcedGrantControls":["Mfa"],
                    "enforcedSessionControls":[],
                    "result":"notApplied"
                },
                {
                    "id":"8ed8d7f7-0a2e-437b-b512-9e47bed562e6",
                    "displayName":"MFA for everyones",
                    "enforcedGrantControls":[],
                    "enforcedSessionControls":[],
                    "result":"notEnabled"
                },
                {
                    "id":"52924e0f-798b-4afd-8c42-49055c7d6395",
                    "displayName":"Device compliant",
                    "enforcedGrantControls":[],
                    "enforcedSessionControls":[],
                    "result":"notEnabled"
                },
             ],
            "isInteractive":true,
            "tokenIssuerType":"AzureAD",
            "authenticationProcessingDetails":[],
            "networkLocationDetails":[],
            "processingTimeInMilliseconds":0,
            "riskDetail":"hidden",
            "riskLevelAggregated":"hidden",
            "riskLevelDuringSignIn":"hidden",
            "riskState":"none",
            "riskEventTypes":[],
            "resourceDisplayName":"windows azure service management api",
            "resourceId":"797f4846-ba00-4fd7-ba43-dac1f8f63013",
            "authenticationMethodsUsed":[]
        }
}
```

A.3 Different Suscriptions of Azure AD

fFree (Included in Azure Sub)

- Limited to 500,000 Directory Objects
- Identity management capabilities and device registration
- Single Sign-On can be assigned to 10 apps per user
- B2B collaboration capabilities (allows you to assign guest users that exist outside of your business)
- Self-service password change (cloud users)
- Connect (syncs on-premise AD to Azure AD)
- Basic security reports

Basic ($1 per user per month)

- Unlimited Directory Objects
- Identity management capabilities and device registration
- Single Sign-On can be assigned to 10 apps per user

- B2B collaboration capabilities (allows you to assign guest users that exist outside of your business)
- Self-service password change (cloud users)
- Connect (syncs on-premise AD to Azure AD)
- Basic security reports
- Group-based access management and provisioning
- Self-service password reset (cloud users)
- Ability to brand logon pages
- Service Level Agreement

Premium P1 ($6 per user per month)

- Unlimited Directory Objects
- Identity management capabilities and device registration
- Single Sign-On can be assigned to unlimited apps per user
- Retrieval of Sign in Logs using Graph API
- B2B collaboration capabilities (allows you to assign guest users that exist outside of your business)
- Self-service password change (cloud users)
- Connect (syncs on-premise AD to Azure AD)
- Advanced reports
- Group-based access management and provisioning
- Self-service password reset (cloud users)
- Ability to brand logon pages
- Service Level Agreement
- Application proxy
- Dynamic groups, group creation, group naming policy, usage guidelines, etc.
- On-premise write back for Self-service reset, change, and unlock

- Two-way sync between on-premise and ADD
- Multi-factor authentication
- Microsoft Identity Manager user CAL
- Cloud App Discovery
- Connect Health
- Conditional Access based on health/location.
- Automatic password rollover (for group accounts)
- Ability to grant conditional access based on location, device state, and group
- Integrations with 3rd party identity governance partners
- ToU
- Sharepoint limited access
- OneDrive for Business (limited access)
- Preview integration for 3rd party MFA partners
- Cloud App Security Integration

Premium P2 $9 per user per month)

- Everything offered in P1
- Identity Protection
- Privileged Identity Management
- Access reviews

Office 365 (Included In Office 365 Subs)

- Everything included in the Free Tier
- Unlimited Directory Objects
- Multi-factor authentication