Jolkkonen Tomi

CLOUD ASSET IDENTIFICATION STRATEGY

UNIVERSITY OF JYVÄSKYLÄ
FACULTY OF INFORMATION TECHNOLOGY
2022

ABSTRACT

Jolkkonen, Tomi
Cloud asset identification strategy
Jyväskylä: University of Jyväskylä, 2022, 63 pp, 1 Appendix
Computer Science, Master’s Thesis
Supervisor(s): Lehto, Martti; Niemelä, Mikko

Securing information is a critical issue, whether you are a private citizen taking pictures and videos, a business owner making backups and using cloud computing, driving a smart car, or using a smart home. Information is collected, used and modified and is ubiquitous. Very often, the solution is to use some sort of cloud service provider.

The common problem, however, is to identify where your data is, and what cybersecurity means are used to keep it safe. The Internet of Things is an ever-expanding world of devices connecting nearly every machine of the world and collecting data from machines and sensors for the cloud. Every business in the world collects and uses data, and servers are often rented from outside the company’s own walls. A private person gathers and uses information every day with a phone or computer using multiple services. The fair question is: Do we know what data we have, where it is and how it stays secure?

The key concern is to identify cloud assets that are used by any operator, private or public.

This thesis presents a cloud asset identification strategy that can be used to fully identify all your assets throughout the supply chain and protect your data. With internet resources and a few open-source tools it is easy to identify organisations’ all domains, IP addresses and which ones belong to cloud service providers. With these steps, an organisation can protect itself better from the vulnerabilities of cloud services. Without knowing what assets you have, it is impossible to protect them.

Keywords: Cloud, IP identification, Asset, Dark Web, Internet of Things, Vulnerabilities, Cybersecurity

TIIVISTELMÄ

Jolkkonen, Tomi
Pilviomaisuuden tunnistusstrategia
Jyväskylä: Jyväskylän yliopisto, 2022, 63 sivua, 1 liite
Tietotekniikka / Sensoriverkot, Pro Gradu -tutkielma
Ohjaajat: Lehto, Martti; Niemelä, Mikko

Tietojen ja datan suojaaminen on kriittinen toimenpide, olitpa sitten yksityishenkilö tallentamassa kuvia ja ajamassa älykästä autoa tai yritys tekemässä varmuuskopioita ja käyttämässä pilvilaskentaa. Tänä päivänä lähes kaikki laitteet ja palvelut keräävät, käyttävät sekä lähettävät tietoa tallentaen sitä useisiin eri kohteisiin. Koska dataa on paljon ja sitä pitää käsitellä monin eri tavoin, ratkaisuna on usein käyttää erilaisia pilvipalvelujen tarjoajia.

Ongelmaksi syntyy, että aina ei tiedetä missä oma data sijaitsee tai mitä keinoja sen turvaamiseksi käytetään. Esineiden internet on jatkuvasti laajeneva laitteiden maailma, joka yhdistää lähes kaikki maailman koneet yhdeksai jättiverkoksi keräten ja käyttäen niiden synnyttämää dataa useissa eri kohteissa, joiden sijainti on hämärän peitossa. Useat eri palvelut myös puhelimissa perustuvat datan käsittelemiseen sekä sensoreiden yhteistoimintaan. Tietoturvan kannalta on siis syytä pohtia, tiedämmekö mitä tietoja meillä on, missä ne sijaitsevat ja kuinka ne turvataan.

Tärkein ratkaisu ja aloituspiste tiedon turvaamiseen on tunnistaa pilviresurssit, joita kaikki laitteet, palvelut ja operaattorit käyttävät. Tämä opinnäytetyö esittelee pilviresurssien tunnistamisstrategian, jonka avulla voidaan identifioidaan henkilön tai yrityksen kaikki domainit ja niihin liittyvät IP-osoitteet ja tätä kautta tiedon sijainnit, jotta niiden suojaamiseksi voidaan tehdä jatkotoimenpiteitä. Pilviomaisuuden tunnistus voidaan tehdä internet-lähteiden sekä avoimen lähdekoodin sovellusten avulla. Tällä strategialla organisaatio voi suojautua paremmin pilvipalvelujen haavoittuvuuksilta. Kaikki lähtee IP-osoitteiden ja pilvipalveluiden tunnistamisesta, koska jos ei tiedetä mitä omaisuutta meillä on, sitä ei voida suojella.

Avainsanat: pilvipalvelut, IP-tunnistus, omaisuus, pimeä verkko, esineiden internet, haavoittuvuudet, kyberturvallisuus

FIGURES

TABLES

TABLE OF CONTENTS

1 INTRODUCTION

The world is heading in a direction where data flows freely in almost all directions. The Internet of Things (IoT) creates new devices and services where computing is done all around, not just at home. Businesses rely on cloud services and private people save their life history to multiple service providers' databases.

One of the big changes due to cost efficiency is that almost all transactions and businesses consist of supply chains. Here lies a new threat to cybersecurity. Organisations often see cybersecurity as a controllable in-house action. Security planning rarely considers third parties, such as vendors or interest groups (Niemelä, 2019). This indicates that companies may not be aware of the amount of information leaking outside their own walls.

This leads to a discussion about whether hackers know this. They are constantly looking for new attack vectors wherever they can easily be found. Hackers are using multiple information sources, from sniffing attacks to dark web forums, from social engineering to social media OSINT. It is safe to reason that hackers prioritise their efforts based on technical vulnerabilities but any other information that is available from a target. The more outsource services or other places for sensitive information are available, the more attacks are going to happen (Niemelä, 2019).

Cloud service providers are an good example of more and more information leaking outside the company's walls. Cloud computing, software as a service, enterprise resource planners (ERP), collaboration tools, frameworks, edge computing and customer relationship management (CRM) systems are often used with the help of cloud service providers (Niemelä, 2019). However, companies using these services may not have a complete understanding of how their data is protected in these services; how authentication and access is controlled, log files and history is kept or deleted, what are the human resources, what is the accident plan, what happens with misconfigurations, what are the policies and procedures. The focus of cybersecurity planning has to be seen as a wider topic than just a

company’s own plans – it has to include the whole supply chain (Niemelä, 2019).

Another point of view is that the future of engineering is on the cloud for two reasons. First, usually cloud service providers have more resources to do computing or other kinds of analyzing than the user. Secondly business always seeks ways to save costs which makes leasing cloud services wise instead of buying new hardware and software yourself (Bermbach, 2021). When cloud services are increasingly popular, there are also more attack surface for the hackers. There may be a whole variety of problems or misconfiguration that happens when two companies - user and cloud service providers - connect. They may use different security protocols, programming languages, scalability and data merge issues. Usually cloud services also cost money, which means that due to cost-efficiency decisions, not all security measures are used (AlMenda, Alzahrani, 2021).

Cloud services are also approaching public sector rapidly. In the future, cloud web services are merging with healthcare beneficial services for human health (Faridi, F. et. al, 2021), education rationalizing their way to manage scarce resources (Sultan, N 2010) not to mention energy sector making it more efficient (Perrons, 2015). Cloud service increase means that data will flow to them from all directions and devices, and this creates new increased need for secure cloud services. We already know what can happen with botnets where devices are used unknowingly for attacking (Feily et. al, 2009). This means that both devices and data may be corrupted which flows into cloud services. A famous example of this was a botnet called Mirai (Sinanovic, Mrdovic 2017).

This thesis concentrates on creating a cloud asset identification strategy for the previous reasons. It is necessary for any person or organisation to understand what cloud assets it has and where they are.

The second chapter goes through cloud vulnerabilities that are found by previous research. The third chapter introduces hypotheses and reasons why it is important to understand what assets you have and why they have to be protected. The main hypothesis is and remains to be that if you do not know what you have, you cannot protect yourself. Another hypothesis is that the dark web is rising and will be a modern way to hide out for criminals and hackers exchanging their information and experiences of vulnerable companies and individuals. This means that there will be an increasing number of attacks directed to not only companies but their whole supply chain. The dark web creates an alluring environment for questionable actions to happen. The third hypothesis is to understand how vast the IoT is. In the near future, not only computers inside the company are vulnerable. Organisations and private people will have computing power all around their physical lives. The IoT changes the whole world in a way that we cannot even imagine. Data will be

flowing from every device, sensor, and infrastructure, and they will be collected to cloud service providers' data centres because it is impossible to have all that data in your own desktop computer or server. It is then safe to assume that the amount of data and cloud services is increasing, which means that it is more important than ever to identify and protect your cloud assets.

Cloud service providers are companies themselves too and they have their own security measures, which may differ from the measures your own organisation is having. By knowing all your cloud assets, you will have a better understanding of how your data is secured by the service provider's own in-house cybersecurity policies. Related to this, it is important to be able to trust your own people as a fifth hypothesis, as well as cloud service providers' staff. By knowing your cloud assets, you will know where people are and how they behave in different situations.

The sixth hypothesis is to understand that we are not talking only about ERP and CRM; we are talking about the whole supply chain in the future, from a guy in the farmland collecting potatoes to the groceries store next door. Everything will be monitored, and data will be collected by machines, devices, sensors and computers. Managing your whole supply chain from manufacturing to production means that you need to control the data flow of all those areas to cloud services. By knowing your cloud assets, this can be done. The seventh hypothesis is to understand that you may already have people you do not know, using your own assets you do not know. In other words, your passwords may already be stolen from your assets or your cloud providers' assets. We have to be able to identify our cloud assets to launch possible proactive or reactive measures to protect the data – or get it back – from the whole supply chain.

In the fourth chapter, a strategy for cloud asset identification is created and tools chosen. As a justification, by creating a working cloud asset identification strategy, an organisation will save money, have a clear understanding of all assets, and have a thorough defence against any attacker.

As a result, new strategy is identified and found useful. Limitation for the study is ever-increasing amount of cloud services all around the world - there is no database containing all cloud services. There are also new services which IP addresses are not recognized as cloud services, for example CloudFlare. We were able to create a strategy for identifying 11 most common services which need constant updating.

More research should be done by better analysing IP addresses' reputations and how to predict attacks from the dark web and clear web against cloud and IoT assets with suitable natural language processing algorithms and machine learning tools.

2 PREVIOUS STUDIES ON CLOUD THREATS

Cloud services encounter several threats against them that are partly the same as threats in any cybersecurity and partly unique only for the cloud environment. The scale of threats is broad and motivations behind them are diverse. There are many angles to look at threats based on the point of view.

Cloud technology is basically servers outside an organisation that can be configured, shared and provisioned to different tasks over the internet. Due to this technology, it has both advantages and risks. Users require understanding what emerging threats, vulnerabilities and countermeasures for them are before they transfer all their computing, data and applications to remote locations (Kaur & Kaur, 2021).

Several new technologies for example smart cities, IoT and 5G need services of cloud computing for processing and storing their information. This means that a wide range of heterogenic new companies will be additional users for cloud services in the future with various different competencies on cybersecurity. Cloud computing involves secure measures of end-users, networks, access management and infrastructures (Kafhali et. al, 2021).

Jain Priya and Jutendra Singh Choihan (2018) did a survey on security issues in cloud computing. They found out that cloud-based services and cloud computing threats and opportunities can be divided into three categories. First, virtualization increases the availability of services for companies, but it also creates a possibility to DDoS attacks. This can be avoided by having a multicloud environment which may then again create other kind of security challenges. Secondly, API-level attacks. Alongside with normal general-purpose clouds like Amazon AWS, there is an increasing amount of specialized cloud services for multiple different needs that usually provide a some sort of API framework for programs to interact together. These services include hosting, device synchronization and streaming. All these services usually expose their own API which may have vulnerabilities. Whatever lock-in kind of problem this may

create, the solution would be to standardize APIs and make compatible software's. The third threat and solution is data confidentiality and auditability in crisis situations. When Sony's PlayStation Network went down in 2011 by an DDoS attack, Sony lost credentials of 77 million accounts. The network was down for days, and millions of players did not know what and when happens next. If a company's infrastructure takes a hit, it should be able to handle the crisis situation. This can be done by deploy encryption, VLANs, Firewalls and different geographical data storages (Priya & Chouhan, 2018).

Another way to divide security threats is by their deployment model. The cloud service providers usually provide three different kinds of layers for the customers. Infrastructure as a Service (IaaS) delivers storage space, processing power and managing organizations databases. Platform as a Service (PaaS) is for organizations that need a particular environment where applications are developed. Software as a Service (SaaS) is providing applications, software, for various needs for organizations. An example of this is ERP (Enterprise Resource Planner) software (Nithiasree et. al, 2021). Further, cloud services are deployed as public clouds (user uses public service that is sold to many other customers too openly), private clouds (user has an exclusive access to cloud infrastructure), community clouds (private cloud is shared with a number of customers) or hybrid clouds.

All these deployment models have slightly different security measurements, but basic security considerations are the same. In organizational security risks, if a cloud service provider goes out of business, this may negatively effect the customers. In physical security risks, the actual location of data is compromised by for example unauthorized on-site access of the data. Technology risks include problems with hardware, resource sharing, portability, software issues and so on (Nithiasree et. al, 2021). The data itself has to be secured too and its privacy, confidentiality, integrity and availability should be checked at all times. Data also has different stages when data is in transit, at rest or used (Nithiasree et. al, 2021). All stages should be included to the security planning.

While doing the security planning for cloud services, a NIST model for cloud computing (National Institute of Standards and Technology, USA) has to be taken into consideration. In the NIST model, there are five different actors in the architecture. A consumer can be an organisation or a person that uses services as well as maintains a business relationship with a cloud service provider. The cloud service provider usually is a supplier of the particular service. An auditor is an organisation that undertakes the evaluation of the cloud service, performance, and the security of the implementation of the cloud. A broker manager the actual use, delivery, and performance of the services and also the relationship between cloud and user. Finally, a carrier is a third party that handles logistic involved in bringing the service to the customer (Shajan &

Rangaswanny, 2021). This means that in a modern multi-cloud environment, there are not only multiple services that need technological surveillance, but a web of people and organisations.

NIST has defined at the basic cloud security requirements that are close to general security measurements in information systems. The requirements are confidentiality (user's data is authorized only to the user), integrity (stored data cannot be tampered with illegally), availability (cloud data must be accessible to the user), privacy (data is used only for its intended purpose), authorization (correct access level provided to the user), authentication (authentic identity of the user and user's data) and accountability (every action made must be established as legitimate by the cloud provider) (Shajan & Rangaswanny, 2021). All these basic security measures are usable in cloud environment and by using them security increases rapidly.

As there are requirements for the safety, there is also a list (made by CSA, Cloud Security Alliance) of top eleven threats for the cloud security. Data breaches involves viewing or stealing protected information without authorization. Misconfiguration and inadequate change control means that in a case where an asset is set up incorrectly, it will leave an asset vulnerable to attacks. Common cause of this is the absence of effective change control, for example unchanged default credentials or disabling standard security controls. Lack of security architecture and strategy in organizations shows in cases where functionality, money and speed are often given priority over the security. Insufficient identity, credential, access and key management means that not strong enough passwords are used, credentials are not protected, or no multifactor authentication is used. Account hijacking is a situation where a malicious actor has access to credentials of other data. Insider threat includes personnel or any people who misuses information from inside the organization; they do n't have to attack anyone, they already have a direct access to data, so it is difficult to defend against the insider. Insecure interfaces and APIs happen when user gets a customized cloud service, but the security is not handled the same way as in some other frameworks. Weak control plane means that while migrating a service to the cloud, sometimes data has to be duplicated or stored to a different place. This secondary momentary situation has to be secured. Metastructure and applistructure failures mean that provider often routinely disclose security operations to protect their systems. How much data must be revealed by the provider is a decision to make, which can also cause misconfigurations. Limited cloud usage visibility creates a situation where an organisation is unable to say if the service is running on their platform is safe or not. And lastly, abuse and nefarious use of cloud services means that an attacker may use cloud resources to target users and this way misuse the cloud resources. This is one way to make phishing attacks, launching DDoS attacks, brute force attacks on stolen credentials and email spam (Shajan & Rangaswanny, 2021). All these attacks have also countermeasures, for example access

management, digital signatures, intrusion detection systems and other security measures for web applications and network. Still, all these threats show how cast and complex multi-cloud environment is to protect and it is clear that not all data nor clouds are protected.

Tahirkheli et. al (2021) made a survey on modern cloud computing security in smart city networks. They identified cloud computing security aspects in three different areas: operations, technology, and management. In operations, main topics to secure a cloud service are put an emphasis on awareness and training, incident and configuration management, contingency planning and maintenance, environmental and media protection, and system information and personnel security. In technology, main topics to concentrate on are access control, system protection, identification and authentication, cloud security audits, identity and key management, physical security protection, backup recovery and archive, core infrastructure and protection, and network protection. In management, a special effort should be put on updated security policy, cloud security strategy process and governance, clear security roles and responsibilities, cloud security guidelines and assessments, service integration, IT and procurement security requirements and cloud security management (Tahirkheli et. al, 2021). Again, many of the topics are common with cybersecurity in general, but there are differences. When we think of the trend of more and more services going to the cloud and usually in the multi-cloud environment, a fair question to ask is whether an organisation is aware of all its assets in all areas. Another question to ask is how many organisations all these topics have covered in their own processes, let alone in their cloud assets.

Cloud services should also be elastic and scalable, easy to use, device and location independent, customizable, and cost-efficient (Shaikh & Sasikumar, 2012). At the same time, there should not be lock-ins and services should always be available without bottlenecks. In a cloud environment full of people, applications, networks, virtualization, identities and data, threats are eminent which calls for a cloud asset identification strategy. First, we need to know what cloud assets we have before we can protect them.

3 KNOW YOUR ASSETS - HYPOTHESES ON CLOUD SECURITY

Eleven years ago, Shatz et al. (2010) introduced a dilemma related to DNA sequencing. Computer performance was followed with a Moore's Law where the speed was doubling every 18–24 months. As a new research area, DNA sequencing needed more computing power than ever, and this created a race between computing and sequencing. The gap was widening and the question of how to design higher-throughput analysis pipelines became crucial. Otherwise, research projects would stall and in the worst-case scenario, even medicine inventions would slow down, and people would be deceased. The gap was closed by inventing algorithms that make better use of a fixed amount of computing power. Unfortunately, these kinds of breakthroughs are impossible to plan or foresee. A more practical option was to develop methods that make better use of multiple computers and processors, thread computing and using other computers for help. This was one of the many reasons why cloud computing became popular (Schatz et al., 2010).

A decade later, everything is in the cloud and continues to be so. The problem is to recognize what data goes to the cloud and from where. It is important to become familiar with owned assets. In this chapter, several hypotheses are introduced as reasons why to recognize assets and why there is a need for a cloud asset identification strategy.

3.1 If you do not know your assets, you cannot protect them

Hackers tend to change their targets based on any information that is available to them (Niemelä, 2019). This means that the more an organisation has assets around the internet, the more information can be available to everyone. Information means IP addresses, social

media, cloud service providers and so on, every credential, vulnerability or address that can be used for attacking or reconnaissance.

When companies plan their security, it is often seen as an externality, and processes or budgets towards a safe working environment are not aligned with the importance of cybersecurity (Niemelä, 2019). The way cybersecurity is measured does not follow the needs of today's security. This means that organisations may not always focus on identifying the amount of information about them that is available for hackers, either in-house or outsourced, including cloud services. Hackers are using any type of information available to them, and companies are not aware of the information they leak out. This way, there is a gap between actual cybersecurity measures completed and actions required, because organisations do not even know what assets they have. If the resources and budgets do not follow the actions, you are not safe.

Forbes wrote an article about securing a multi-cloud environment. The main point was that there is a good chance that every company nowadays operate in a multi-cloud environment. It offers advantages, for example flexibility for best services (Forbes, 2021). The challenge comes when it is needed to secure owned data. The person who is selling a service for the organization, may himself sell a package that contains several parts from several different providers. Maybe the sales rep does not know either (Forbes, 2021). When choosing a vendor, it it important also to discuss about the types of security measures. This leads to first hypothesis:

If you do not know your assets, it is impossible to protect yourself.

3.2 The dark web is rising

The dark web is a part of the internet that is not indexed by search engines and requires a special browser created for that purpose, which recognises for example onion-ending pages. One of these browsers is called the Tor browser (Dingledine, 2004). Today, the dark web is no longer an invention, and it has already crossed the news threshold of mainstream media. It is an ever-expanding part of the internet from which many of today's attacks rise together with the clear web (normal accessible internet sites). The academic world is exploring the dark web more, although there is still too little research on how marketplaces work, anonymous cryptocurrencies move and how the mind of an attacker fluctuates. Not all criminals use the dark web at this stage, but the number is growing exponentially (Figure 1).

Figure 1 The dark web grows exponentially (UN Office, 2020).

The dark web has many definitions that have changed since its inception. The internet is commonly divided into three categories: clear web, deep web, and dark web. The clear web is the internet that everyone can see or access, e.g. through Google Search Engine. The deep web is the part of the internet behind usernames and passwords, and it is not indexed by search engines. The dark web is part of the internet that is intentionally concealed from the public eye (Finklea, 2017).

The dark web – and more precisely, Tor browser – was designed in 2004 to create a new low-latency communication service between people (Dingledine, 2004). It had no bad purposes in the beginning; creators wanted to create something with better secrecy, congestion control, directory servers, integrity checking, configuration exit policies and practical design. Instead of using addresses that end with .com or .fi, in Tor, all addresses have an end suffix called .onion.

The main idea behind the dark web is anonymity. Users are masked in a way that their internet traffic goes through several constantly changing servers. This way, no one knows who the user is and where he comes from (Dingledine, 2004), which creates an environment for volatile information. Because of its anonymous nature, discussion groups for hackers, dark marketplaces with questionable items for sale, cryptocurrency (an anonymous money) and secure emails or chats are common on the dark web (Finklea, 2017). The dark web acts as a forum for conversation, coordination and action for criminals, terrorists and other malicious actors. There

are, however, normal people too who feel brave enough to try something a little bit different.

If we take a look at the dark web as a customer or a user who has an appeal towards criminal activities, there are differences compared to real-world crimes. Guitton (2013) researched the available content in Tor's hidden services. The results show that marketplaces and discussion forums have differentiated themselves based on content (Guitton, 2013). Black markets are another forum group that has the whole variety of illegal items to purchase via bitcoins (Guitton, 2013). The dark web is a separate corner to the internet, where marginal topic discussions are established in all variety of topics.

Villalva et al. (2018) compared how the use of leaked account credentials differs in the Dark and Clear (or Surface) Web. They researched this topic by setting up honeypots to allure criminals for action. The results were to find four types of attackers from the dark web. First, Curious Ones log into the honey accounts, but do not perform any further actions in them. Second, there are gold diggers who perform searches on the emails contained in the account to find sensitive information that can be quickly monetised. The third group are spammers who use honey accounts to send spam and exploit the trust that contacts have with the account owner. The fourth group are hijackers who change the account password to take full control of it, preventing the original owner from accessing his or her own account (Van Hout, 2013).

On the dark web, so-called pastebins are sites that are created for many purposes, for illegal content. After they are created, no one monitors those sites. A pastebin site can exist for nine minutes or three months and consist of sensitive information about basically anything. This means that stolen items are more vulnerable on the dark web, and they stay there for a longer time (Van Hout, 2013).

This means that the dark web interests many kinds of people, good, normal people, and the dark environment can allure them to do questionable things. A hacker can do dozens of different things, and one of the human-related attacks is phishing. In phishing, an attacker lures the target user to give away critical information (Neshenko, 2019). These credentials are then either shared, sold or used in different places, and naturally the dark web is an important platform for this. It is important to understand that not all hackers are high-end professionals, so after a phishing attack, leaking information can be either intentional or by accident (Neshenko, 2019).

The dark web forums are full of information about social engineering and how individual characteristics guide peoples' behaviour, not to mention the big five personality traits (neuroticism, extraversion, openness to experience, agreeableness, conscientiousness), which are powerful tools for human-hacking like phishing (Warkenting et. Al, 2012).

Another danger is that the dark web is alluring for the average citizen. There is a theory of nice people doing questionable things questionable (Masson, Bancroft, 2017) where they found that the drug marketplace is not really just a marketplace but a place where similar-minded people can exchange their thoughts and experiences of drugs, for example, and even change their drug to a stronger one. This is a logical example of nice people who start to behave in a questionable way when an opportunity is there. It may be seen in statistics where people's attitude towards the dark web is changing to the more positive side as they do not see themselves as criminals, even though they do those kinds of things. (Masson, Bancroft, 2017).

Nowadays you do not have to be an elite hacker; you can buy all the tools and credentials even though you do not have the programming skills (Kwon et al., 2017). Not all people are bad on the dark web, as mentioned before. Jardine (2018) studied privacy, censorship, data breaches, and internet freedom and the drivers of support and opposition to dark web technologies. The interesting result was that most people are opposed to dark web technologies whether they are used for noble and nefarious purposes (Jardine, 2018). Also, exposure to online crime did not get as high results as losing privacy or censorship. This means that the threshold is low when doing questionable things on the dark web.

Crime is also cheap on the dark web. Recently, journalists described the prices for principal hacking services that it is possible to acquire online. According to Business Insider, an individual who wants to hack someone’s Gmail account will have to pay about $90. Hackers, for example, could be hired to hack into a social media account, and the cost to hack into someone’s Facebook account is $350. The investigation conducted by the journalists revealed that a hacker can compromise a Netflix account for just $1.25. Other common commodities in the hacking underground are the hacking courses that can be bought with $20 (INFOSEC, 2021).

In the dark web marketplaces, most users continue their trading activity in a single coexisting marketplace, typically the one with the highest trading volume. User migration is swift, and the trading volumes of migrating users recover quickly. Although individual marketplaces might be closed down, coordinated user migration happens swiftly (Elbahrawy, 2020). This means that if you lose your cloud asset information and it becomes for sale in the dark web marketplace, it may be difficult to get it away from there because marketplaces usually change over time.

Another thing that motivates the criminal is the dark web’s variety of services. You can find marketplaces, file sharing, general discussion about any topic, forums, education and training, information sharing in general and criminal connectivity (Dalins, 2018). New innovative drug markets lure people to run by these anonymous cryptocurrencies and retailers using anonymising technologies. These markets are international, and they were based

on the reputation system. If you sell something, and you cheat the customer, your reputation as a cheater will spread fast and soon no one would buy anything from you. It is to a criminal retailer's advantage to be a trustworthy seller to stay in business. In this strange way, black markets work even better than marketplaces in the clear web, where it is normal to get something strange products from the Chinese Web stores or get cheated in advertising. This is innovative retail in which, in a strange way, we could learn from these criminals. These markets are shipping all around the world and 30 percent of shipping comes from the USA (Aldridge, 2015).

This new innovative dark corner of the internet naturally invites innovative hackers in. If you look at the HackerOne 2021 Hacker Report to understand even white hacker motivations, you can see that 85% of hackers want to learn and advance their career, 76% want to make money and 65% want to have fun. This behaviour is very typical for any programmer or hacker, white or black hat; they are the first ones to say that they are lazy and curious. They want to find a way to do things easier and get paid. This is where the dark web – once again – serves this purpose. You can do everything from home, make money, get achievements, stay anonymous, and decide your amount of effort. It is very alluring to control your own destiny completely.

This lengthy introduction to the dark web underlines the importance of knowing all your cloud assets. There is a new corner in the internet which consist of an environment that can make a normal people behave in a non-normal way (for example cloud service providers employees), everything is anonymous, credentials and personal information is sold with anonymous money, all kinds of forums, tools and help is available. All of these multiple examples lead to the second hypothesis:

By knowing your assets, you will get a better defence against ever-increasing criminal activities on the dark web.

3.3 The exponentially growing IoT

One of the biggest and still a bit newer aspects of information technology is the IoT. This refers to all devices that are not computers but have so-called intelligence, meaning that they can be connected to the internet or other devices on the internet (IBM, 2021). Equally, we are seeing exponential growth in IoT devices and applications (Vipinraj, 2001), linking the physical world to information technology (See Figure 2). It is only a matter of time before IoT is using the majority of the cloud service providers.

When the whole physical world is connected to the internet and every device becomes a computer, the game is completely different

from a security point of view. There is reputational damage, loss of customers and distrust of IoT devices, even in the case of medical devices (Moor & Anderson, 2018). 48% of device owners are not aware that their device can be attacked, and 40% never have updated their IoT device firmware (Moor & Anderson, 2018).

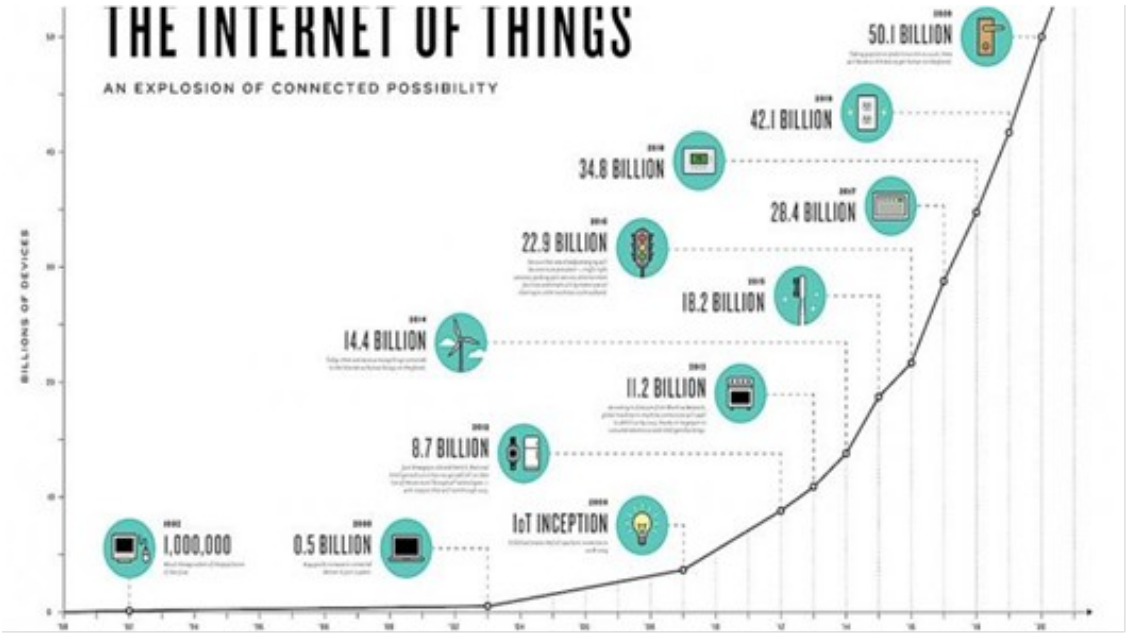

Figure 2 The exponentially growing IoT (NCTA, 2014).

Soon not only computers are connected to the internet, but every device will have a connected computer in them, from washing machines, factories and hospitals to cars (Moor & Anderson, 2018). This naturally means that all those things will save their data somewhere, which will be in the cloud. The IoT is a collection of distinctive heterogeneous devices with various kinds of technology with insufficient security measures in them. Security measures are not as high as in more powerful central computers and servers, which means that microcontrollers, communication protocols and sensors are more vulnerable to attacks. This means that the data they are creating and sending to the cloud are more vulnerable.

There is a danger that people and organisations do not understand what kind of assets they have in the future, where sensor networks and IoT devices are everywhere. Earlier this year, a fish-tank thermometer was used to make a cyberattack (Marks, 2021). Another example is to search www.shodan.io to see how many refrigerators are open to the internet at this very moment (meaning they have open ports for attack) (www.shodan.io). The number is vast. The third example is to understand how the IoT and Sensor Networks work. Information is not only transferred straight to the database, which could sometimes be safe and guarded by a cloud service provider. Sensor data may be altered, pre-filtered or modified beforehand, in the sensor network sensors, sinks or microcontrollers, before they reach the safety of the database (edge computing). In other words,

data are not vulnerable only in the beginning of the whole supply chain, or in the end part of the chain in the cloud, but in any part in the middle. This ‘raw data’ or somewhat altered process data can leak straight to the attacker, or this modification data may be saved by a cloud service provider you do not know exists. This leads to the third hypothesis:

In the near future, there are an increasing amount of data and devices in the physical world, and these are all assets to be aware of.

3.4 Can you affect your cloud service provider’s cybersecurity measures?

Organisations tend to concentrate on making their own processes and servers secure. They use a lot of resources and money to take care of the most powerful tools and firewalls, virus scanners and staff training to be safe. One thing that may be forgotten is that organisations and their third-party members use dozens of cloud services all around the internet and they do not have any control over the security measures that cloud service providers use (Niemelä, 2016). Even though the security measures are done, they may differ from the measures you have taken. If the cloud provider is a big international company, it may not be ready to share all the security information with a single customer. This may lead to at least misconfigurations of your data or different kinds of preparations against crisis situations.

Niemelä and Koistinen (2020), in their book Smiling security, talked about internal breaches and other organisational issues critical to security: the whole supply chain should be secured. Data migration to the cloud should be checked because when data is sent to the cloud, individual data security controls no longer work. How to monitor and protect owned data with a service provider that does not let to do all the things an organisation wants, is a big question. Data liability in the cloud is important: the more data exists, the more can be leaked. The knowledge of all the data you have and where it is, with history logs, is important. The executive-level operators should make sure that the budget, resources, support, and communication are working, not to mention risk management and disaster plans. The key personnel, both in-house and in-cloud service providers’ staff, should know what facilities there are against physical breaches, are the machines themselves in working order, what kind of computers they have and are they updated, or how access, authorisation and authentication is controlled. A nice-to-have information would be the history of data breaches with your cloud provider.

Kyberturvallisuuskeskus (Cyber Security Center in Finland) has published criteria to assess the information security of cloud services

(Traficom, 2019). The manual consists of several steps to make sure how to know your cloud service providers' security means and make this way your own data more secure. These steps are for example risk assessment, service and deployment models, location, framework conditions, management, personnel, physical security, information system security, data security, operations, transferability and compatibility and change management (Traficom, 2019). This leads to the fourth hypothesis:

By knowing your assets, you know your cloud service provider assets and this way you control your own data.

3.5 Trusting people

Technology is not the only security threat in cloud assets as in any other assets either. Another one is people. It is a valuable information to know how people are hired in cloud service providers - internal breaches are a big problem (Niemelä & Koistinen, 2020). . Even though everything is secured from the outside world, there are fewer prevention tools if there is an in-house breach.

By taking a closer look at the personalities and characteristics of different people using the internet, there are four major factors that differentiate between internet and face-to-face action: greater anonymity, the diminution of the importance of physical appearance, greater control over the time and pace of interactions, and the ease of finding similar others (Amichai-Hamburger, 2009). There are similarities in the technology of the dark web. The dark web's biggest advantage is to stay anonymous, so the importance of physical appearance is not important, and this makes a person braver. When someone has bad intentions, this person will find people with the same interests on discussion forums. Based on Erik Erikson's theories, the internet can even help people develop a sense of coherent identity. Also, based on another theory created by Albert Bandura, help may be given to others over the internet at a low cost (Amichai-Hamburger, 2009). This explains why discussion forums are so powerful places to spread information about everything, even bad habits.

Amichai-Hamburger (2009) explains Erich Fromm's theory about five basic human needs: a need to relate to others, a need for transcendence, a need to be rooted, a need for identity and a need for a frame of reference. If a person lacks some of his or her needs in real life, he or she may want to find them either from the clear or the dark web. This way it is possible that if a person has problems in either personal life or at work, this person may be guided to blackhat hacking, terrorism, or espionage – not because he wants to do harm but because he wants to be a part of any group, to consider that his

new family outside real world. There is a need for closure, cognition, feeling of control, sensation-seeking and risk-taking that can sometimes explain risky behaviour (Amichai-Hamburger, 2009). All these characters are in favour of the thought that the dark web – and internet in general – encourages normal people to do questionable things, which leads to the conclusion that people are also a threat to cybersecurity. And because human beings are assets, and there are human beings working in cloud service providers' offices, this is an asset that should be recognised.

Dolliver and Kenney (2016) further researched drug vendors in Tor Networks (the dark web) and their characteristics. They found out that buyers are using the internet because you can do it from the comfort of your own home without engaging in face-to-face open-air communication (Dolliver & Kennedy, 2016). There may be cultural differences between people's risk-taking options because drugs were usually sold from the USA, the UK, Germany, and Australia, when other items were sold globally (Dolliver & Kennedy, 2016). Nevertheless, during the covid-time when people are at home, this is one more allure towards unhappy people behaving in unnormal way.

People may have many kinds of motivations to use someone's credentials or sensitive information for personal benefit. These motivational goals can be listed as ten value types: universalism (appreciation of all kinds of people), benevolence (preserving the welfare of everyone), conformity (self-discipline), tradition (hacker culture), security (anonymity), power (access to tools and people), achievement (praise among your kind), hedonism (personal success), stimulation (challenge in life) and self-direction (it is in your own hands) (Madarie, 2017). Although motivation and hacking activities are not always straightforward, there are similarities. Intellectual challenge and curiosity seem to be big factors. There are people who do not have clear motivation for their actions and are acting based on their gut feelings (Madarie, 2017). Based on this, even random acts of carelessness can become virtual attacks against any cloud or normal asset.

Not only people, but people groups and countries use the dark web. Espionage has taken a big leap from face-to-face information gathering to global computer-to-computer web harvesting. As long as we have had secrets, there has been espionage (Merritt & Mullins 2011).

Human beings are the target of phishing attacks, which are part of social engineering. This is yet another reason why people are an important asset inhouse and in cloud service provider's staff to know well. This non-technical strategy, which uses psychological manipulation and persuasive communication to deceive users into making security mistakes or giving away their credentials, such as passwords, bank information, access to systems, or money, is an attack vector in almost every major cyberattack. The art of exploiting human psychology to gain access to organisations' systems, buildings

or data. Social engineering attacks, unlike hacking, do not use technical expertise because attackers rely on social psychology. Statistics reveal that 52% of breaches featured hacking, 28% involved malware, and 32–33% included phishing or social engineering. 92% of malware is delivered by email and 34% of data breaches involve internal actors (Tamber, 2021). This means that almost every type of cybersecurity attack involves social engineering. The FBI (Figure 3) investigated crime types in cyberattacks and phishing is the number one attack form and it is purely based on social engineering.

2019 CRIME TYPES

By Victim Count

Crime Type	Victims	Crime Type	Victims
Phishing/Vishing/Smishing/Pharming	114,702	Lottery/Sweepstakes/Inheritance	7,767
Non-Payment/Non-Delivery	61,832	Misrepresentation	5,975
Extortion	43,101	Investment	3,999
Personal Data Breach	38,218	IPR/Copyright and Counterfeit	3,892
Spoofing	25,789	Malware/Scareware/Virus	2,373
BEC/EAC	23,775	Ransomware	2,047
Confidence Fraud/Romance	19,473	Corporate Data Breach	1,795
Identity Theft	16,053	Denial of Service/TDoS	1,353
Harassment/Threats of Violence	15,502	Crimes Against Children	1,312
Overpayment	15,395	Re-shipping	929
Advanced Fee	14,607	Civil Matter	908
Employment	14,493	Health Care Related	657
Credit Card Fraud	14,378	Charity	407
Government Impersonation	13,873	Gambling	262
Tech Support	13,633	Terrorism	61
Real Estate/Rental	11,677	Hacktivist	39
Other	10,842		

Figure 3 Crime types in cybercrimes (FBI, 2019).

Whatever the reason behind the questionable actions of human beings, it is clear that it is an important asset to recognise and identify, both in your own organisation and in cloud services. The internet, especially the dark web, is a place where tools and people can connect, whether you use them to find yourself a virtual family, just have fun or for curiosity. This leads to the fifth hypothesis:

Human resources are also your assets. By knowing your cloud provider, you also need to know who works with your credentials.

3.6 Managing the whole supply chain from manufacturing to production

In many international or listed companies, manufacturing and production processes are heavily dependent on partners, e.g. a supply chain. This means that hackers may be interested in doing a reconnaissance on which partners are the most important ones, or who are the people behind key information (companies' top clients and suppliers) (Niemelä, 2016). This information can also be drawn from social media like Twitter, Facebook, Instagram, LinkedIn and so on. By connecting information, new knowledge is created with OSINT techniques that are not safe behind firewalls. With all this information, a so-called asset discovery can be done (blogs, web pages, online surveys, marketing materials, extranets, intranets, file transfer terminals, Facebook fan sites, LinkedIn company pages, campaign sites, job ads, etc.). Telecommuting is also a new trend in the world, where everyone is working outside the office with VNC connections. Even though passwords are often tricky to break, backup copies are often unencrypted, and they are in the cloud (Niemelä, 2016). Companies also tend to take backups regularly, with all unnecessary information and mistakes in those backups.

These are all ways for a hacker to get into a company's systems, and backups with all the history are almost always in the cloud. Not to mention open ports and the sea of Wi-Fi. For some reason, the oldest computers often control the most critical systems and are still used to surf the internet. Also, the person who lifts the potatoes from the field and has sensors in his body for cost-efficient work in the future will not have the same security procedures as the organisation's IT department. The whole supply chain creates information in the future, and it can be long and worldwide.

Accenture has published a report about securing supply chain. Key ideas are that as supply chains become more complex and connected, almost half of the cyberattacks are originated in the extended supply chain, not the enterprise itself. The report suggests that one of the main fixes is to enhance the visibility across the entire supply chain network (Accenture, 2020). This leads to the sixth hypothesis:

Knowing your assets means that you know, defend, and control your whole supply chain.

3.7 Someone you do not know, already has your assets you do not know

It is also possible that your passwords are already stolen. This means that all security procedures are late and there are assets somewhere you do not know you have, and they are accessed regularly by someone unknown. It can be someone in the farmland who is collecting potatoes for a potato company and who uses sensor networks to enhance the work. It can be a communication protocol over the Atlantic Ocean that uses multiple satellites. Or it can be the groceries store or a vendor at the end of the supply chain.

An example of a situation where data leak is unknown is so called Man-in-the-middle attack. In it, an attacker positions itself in a conversation between a user and an application - in this case between an organization and a cloud service provider. One of the most famous ones was revealed by Edward Snowden when NSA disguised itself as Google to spy (Moyer, 2013).

One thing is to know your assets; another thing is to know how many people have the possibility to access your information. It is impossible to protect yourself if you do not know your assets or your people. This leads to the seventh hypothesis:

Someone you do not know already has your assets.

4 DATA AND METHODS - CREATING A STRATEGY FOR CLOUD ASSET IDENFITICATION

4.1 Premise of the research and research question

As the hypotheses and previous studies on cloud threats show, there are many reasons to create a strategy for identifying our own assets, especially in the cloud. Based on these challenges, the main research question is:

What means are needed for an organisation to use cloud services in a safe way.

As a result, we will have a strategy for cloud asset identification. This strategy is made by choosing an organisation, finding its domains, recognising IP addresses within these domains, identifying cloud assets from those IP addresses, calculating the cloud asset percentage, and creating a cloud asset strategy. All steps are important and require different set or tasks and tools.

4.2 Research Method and Strategy

In information systems, we can categorize research approached based on what kind of results we get. One way from many ways is to divide research into theory and empirical study. An empirical study includes a case study that is used in this research, because it can be used when research and theory are at earliy stage (Roethlisberger, 1977). Case studies can then again be divided into qualitative and quantitative approaches. As any research method has its advantages and problems, I have chosen to use case study and qualitative

method in this research. Also in case study, research can study phenomena in its natural setting and learn about the state-of-the-art things and generate new ways to handle new things. The case method allows to understand the nature of the new problem in hand. Case study is a good choice when there is not much research done in that research question (Benbasat et. al, 1987).

There is a lot of research on cloud vulnerabilities, but not so much on IP recognition of cloud services, which means that there is not quantitative data or previous research on the topic. For these reasons, I have chosen case study as a research strategy.

4.3 Research process

4.3.1 Data Collection and platform

The strategy is to recognise all domains and IP addresses within an organisation with a set of tools, analyse the received data and react to it by creating a strategy for cloud assets. Data (here domains and IP addresses) can be collected from public web-based services and internet sites, also open-source tools that are available for free to use. As a platform, any computer can be used with an internet connection. In this research Linux Ubuntu PC computer was used, but the same research can be created with any operating system.

4.3.2 Ethical side note and reliability of the research

To be impartial and use only tools and services that have no business preference, the chosen tools are open-source tools or otherwise neutral and free to use. There were over 100 tools that were filtered out, both commercial and free ones, with only criteria for IP recognition and internet traffic analysis. Other options were available, but the goal was only to identify IP addresses and cloud service providers. The tools were searched from the internet and mainly Stack Overflow, Github, Gitlab and other similar sites. Different kinds of software review sites were used to recognise the most suitable tools for packet tracing, vulnerability scanning and IP recognition. Many more tools may be available for various purposes, but these were chosen based on the research question. The whole list is in Appendix A. The purpose of the research was only to create a research process for identifying cloud assets from IP chosen IP addresses.

4.3.3 Comparing tools

Before choosing and using correct tools, an important task was to compare open-source tools related to internet traffic analysis. It started with finding any tools that could be using the IP information of any service, so over 100 different tools were screened related to network traffic identification, cybersecurity and pen-testing tools and so on (Appendix 1). Both commercial and free/demo tools were tested, but it was quickly realised that there are lots of tools that are not suitable for this task. It was decided to concentrate only on easy-to-use open-source tools instead of going through the whole list of over 100 different tools with dozens of different options other than IP identification. Finally, a few tools were chosen and compared: IP2provider, Shodan cli, which-cloud, cloud_ip_ranges and server-ip-addresses.

Which-cloud is a six-year-old open-source tool that is easy to use. So even though it recognises IP addresses, whether they belong to the cloud or not, the cloud database is old. It could not search for more than one address at a time. After editing the code, it was too complex for this solution.

Shodan CLI can be used to identify whether an IP address belongs to the cloud or CloudFlare (cdn). This option is also free to use without an account upgrade. It has a search option called IP lookup, where you see right away if a certain IP address belongs to the cloud, CloudFlare or on rack. Shodan is not open source, but it is free to use up to some point of using API upgrades, and the IP lookup option does not require any bought account memberships, although new options may be available that way.

Cloud_ip_ranges works like which-cloud; it can be used by searching single IP addresses, and it should be updated somewhat regularly. Then, a Docker/YAML file called server-ip-addresses can be used to get updated lists of IP ranges. This may work well with which-cloud and Cloud_ip_ranges if a third option is needed besides IP2provider and Shodan, which are the two best options.

IP2provider is simple to use, and it is updated regularly by newest IP ranges. It also has a piping option to search multiple IP addresses at once. It updates all cloud service providers' IP ranges every week, and you can update the cloud data to the program with a single command. Piping commands works in the terminal, which makes it easier to have multiple IP searches at once. This is done by creating a .txt file from IP addresses or any domain and piping it with the program itself. It prints out every IP address that has a cloud in it and also which cloud provider it is.

Server-ip-addresses is a resource for updated cloud server IP addresses.

4.3.4 Choosing two best tools

A comparison between IP2Provider and Shodan was done in the end. Previously listed tools were filtered into these two options:

IP2Provider

An open-source tool to check which cloud provider is hosting a particular IP address. Some providers will also have service and region listed. This is done by a user called ‘oldrho’. It is done using Python, and it is run simply by giving the IP address as an argument to the program: ./ip2provider.py 52.4.0.0

It can be piped by creating a .txt file of all searched IP addresses and piping it to the program: cat ip_addresses.txt | ./ip2provider.py

Supported Cloud Service Providers: Amazon Web Services, Microsoft Azure (Public and Government Clouds), Google Cloud Platform, IBM/SoftLayer Cloud, Oracle Cloud, Alibaba Cloud, Linode, DigitalOcean, RackSpace, Cloudspace. It also recognises CloudFlare, although it is not clear how.

Shodan CLI

Shodan is the world’s first search engine for internet-connected devices. It has both free and payable account versions, as well as both Web UI and CLI versions. An example of how to see information about the host where it is located, what ports are optn and which organisation owns the IP: shodan host 52.4.0.0

Search examples: search fridges:

- shodan search --fields ip_str,port,org refrigerator
- shodan download fridge-data refrigerator
- shodan parse --fields ip_str,port,org fridge-data.json.gz

The tool that was chosen was IP2provider, with support from Shodan:

- It is fast and simple to use
- It is open source
- It recognises 11 biggest cloud service providers

Limitation: you need to update it manually, although IP ranges are updated by IP2provider:

- More cloud service providers require code updates
- With these checked clouds: Amazon Web Services, Azure, Google, IBM, Oracle, Alibaba, Linode, DigitalOcean, RackSpace, Cloudspace, CloudFlare

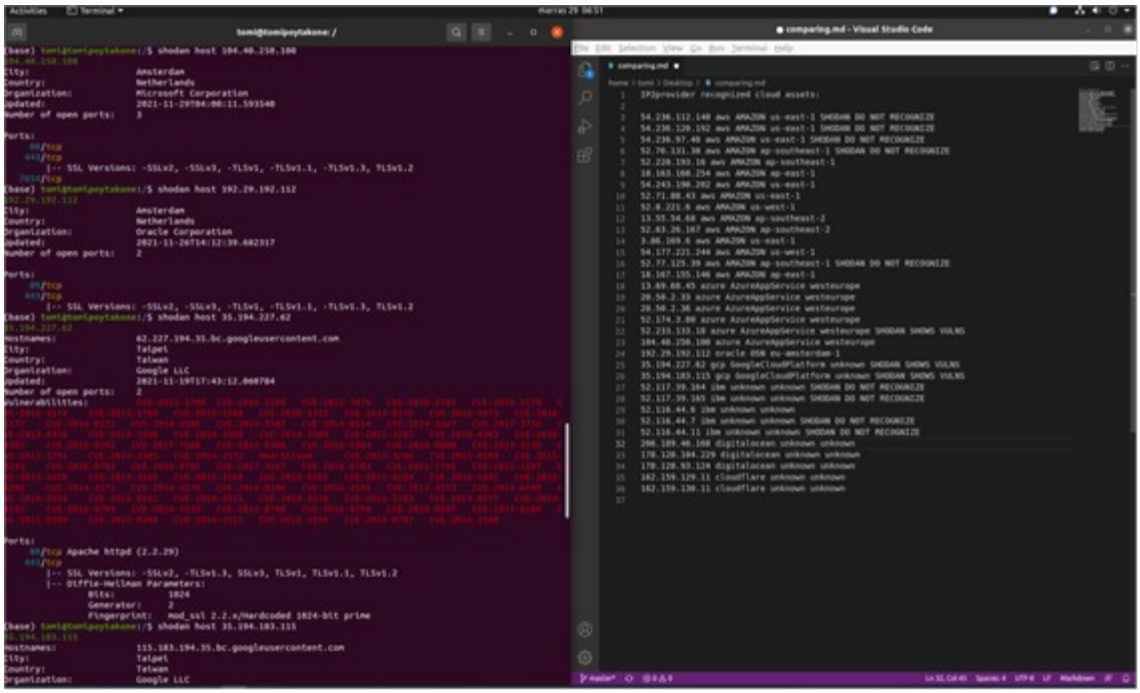

Figure 4 Comparison of Shodan and IP2provider.

4.3.5 The research process - Cloud asset identification strategy

The first task was to choose an organisation that needed a cloud asset identification strategy. Ten different companies were chosen based on their Cyber Exposure Index (www.cyberexposureindex.com) rating as a sample.

Every company was chosen from a different country, a different industry, and a high enough exposure score to underline the importance of knowing cloud assets and have a variety between results. Any other way of choosing an organisation is also suitable.

The sample list of companies were:

- CSL Limited (Australia, Healthcare)
- Stora Enso (Finland, Materials)
- Grenke AG (Germany, Financials)
- Sojitz Corporation (Hong Kong, Industrials)
- PT Ace Hardware (Indonesia, Consumer Discretionary)
- BioDue S.p.A. (Italy, Consumer Staples)
- Ellies Holding Limited (South Africa, Information Technology)
- Safestore Holdings plc (United Kingdom, Real Estate)
- HollyFrontier Corporation (United States, Energy), and
- Hyflux Ltd (Singapore, Utilities).

The second step was to find out what domains any organisation has. There are several ways to find domains owned by a company from the internet. In this research, a website Cyber Exposure Index site was again used. The web site shows all domains linked to a certain

listed company. Again, it is not important what service to use for getting domains of a certain company, any register is suitable.

The third step is to find out which IP addresses are shared or linked with these domains. The Linux command line can be used to recognise IP addresses and domains (whois, dig, nslookup, ping), also www.shodan.io is a web site that has information about both IP addresses and domains. Cross checking is needed, because every domain may have differentiating IP addresses, and every IP address may have moving domains and sister / shared domains. One tool that can be used to obtain basic information about all linked domains is Robtex (Figure 5, www.robtex.com). This part takes a lot of time without ready-used tools; a consultant could easily charge thousands of euros from this part only. For this research, I had access to the Cyber Intelligence House (https://cyberintelligencehouse.com) platform, which shows all IP addresses linked to searched domains.

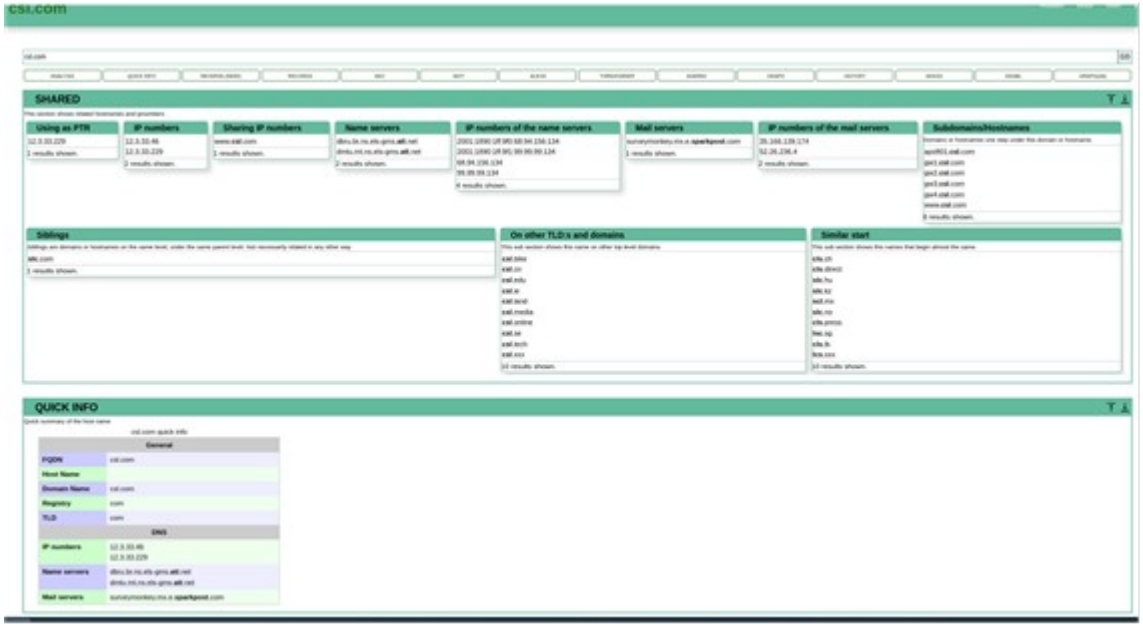

Figure 5 Robtex tool to recognise domains.

The next step was to create IP address list of each company and their domains. A simple excel file was created. The IP address list was then used with two different open-source tools to test which IP addresses belong to cloud service providers and which ones are their own servers. These tools are called IP2provider and Shodan CLI (Figures 6 & 7).

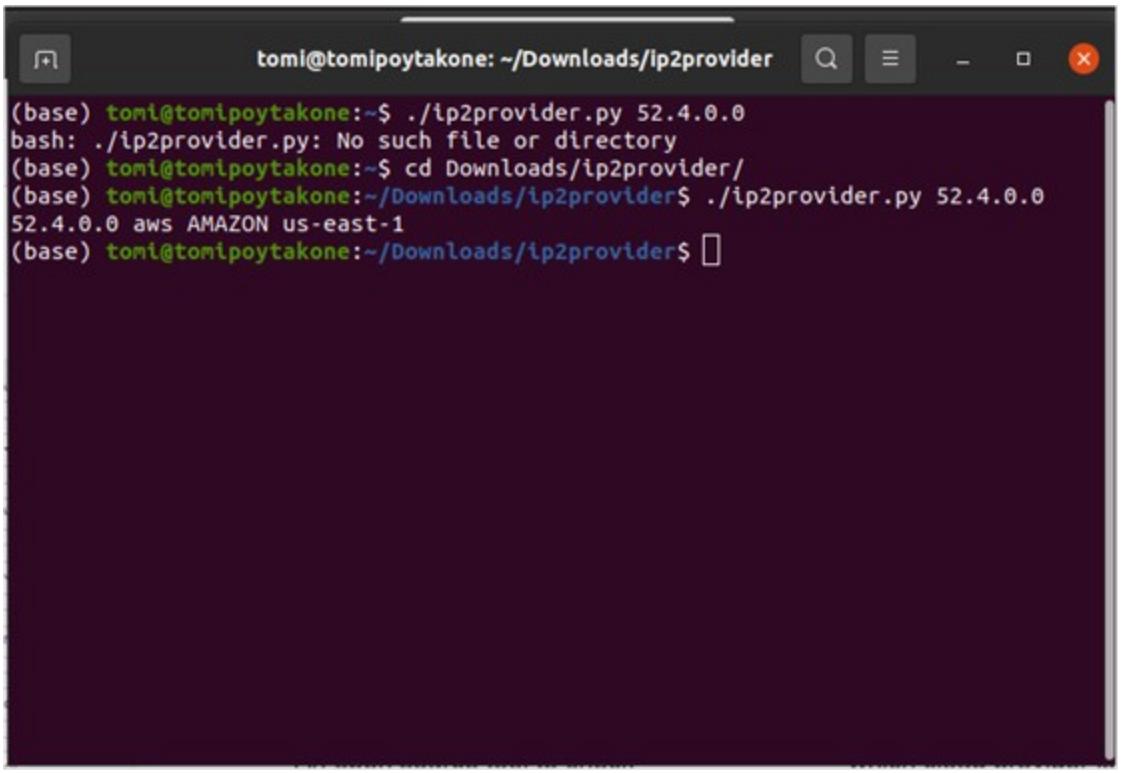

Figure 6 IP2provider usage.

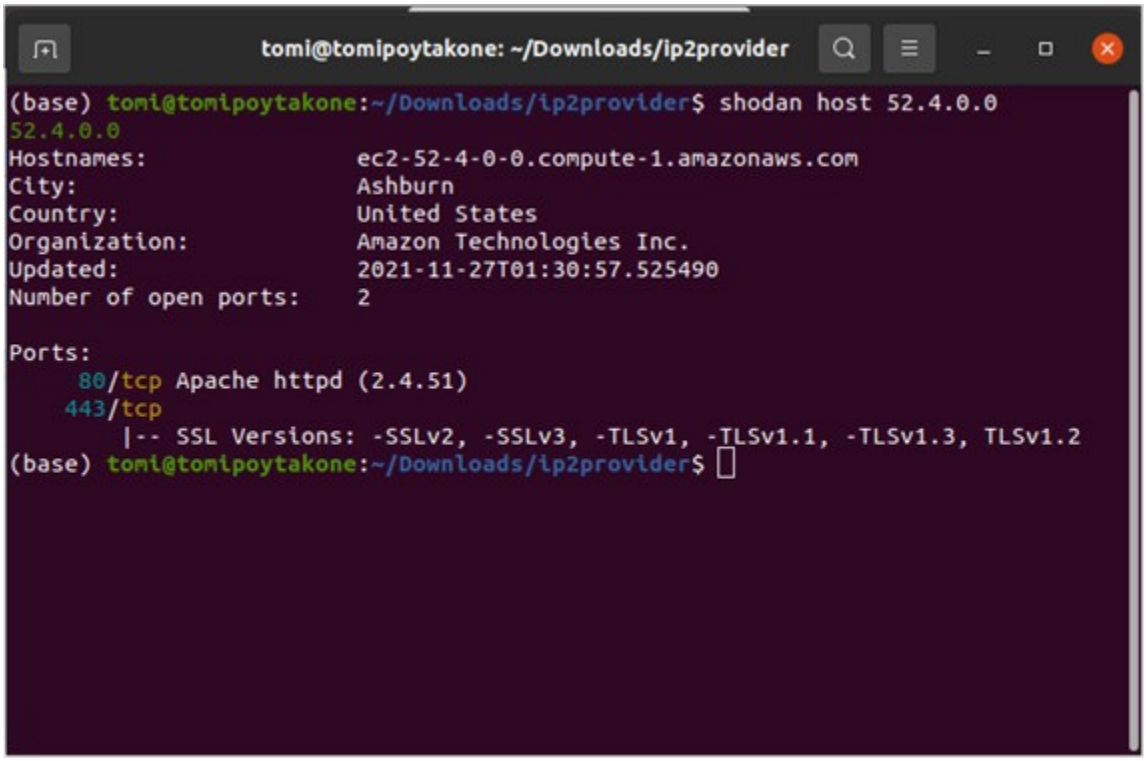

Figure 7 Shodan Command-Line Interface.

After all IP addresses were recognised as either cloud or normal IP addresses, cloud asset percentage was calculated to understand how much cloud assets a certain organisation has (Figure 8).

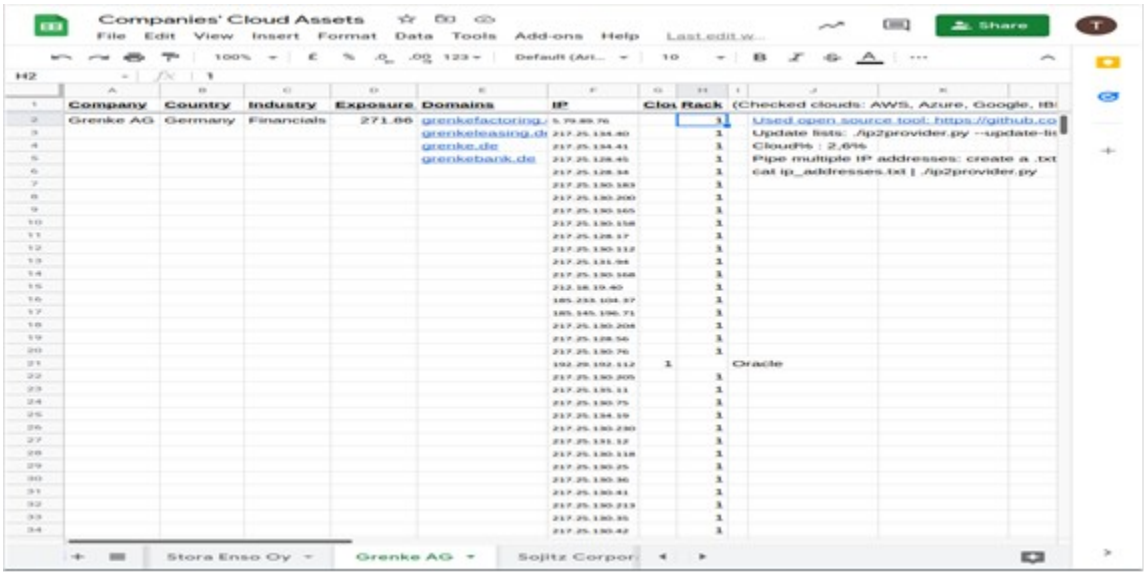

Figure 8 Identifying cloud assets.

In the end, we identified IP cloud assets for ten companies from ten different countries and ten different industries. Cloud percentages varied quite a lot between companies; some companies did not have any cloud assets, and some companies had a cloud asset percentage of 17.5%. Based on this percentage, a different strategy for each company should be created, based on the hypotheses shown in Chapter 3.

Different cloud service providers usually sell products for different purposes (ERP, CRM, database, cloud computing, managed services, etc.), which means that a ‘cloud map’ should be drawn for each organisation with a more detailed plan on how to get security measures and data flows recognised for all cloud providers. Figure 9 shows the entire cloud asset identification strategy process.

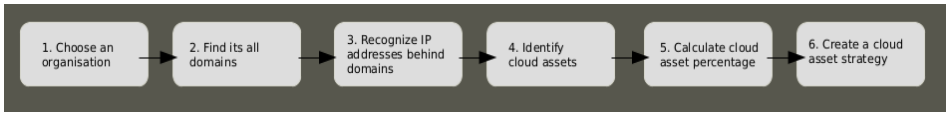

Figure 9 Cloud Asset Identification Strategy Process.

4.3.6 The scope of the research

The scope of this thesis is to concentrate on cloud asset identification. There are many reasons also to identify every kind of asset an organisation has, as Chapter 2 and 3 shows. IP addresses can also be used for ‘reputation’ recognition (OSINTme, 2021). In phishing attacks, fake IP addresses are used against users. The malicious actors create hundreds of fake domains, impersonating banks, post offices and any other services that people use. By recognising IP addresses based on their listed reputations, these kinds of attacks can be defended against. The IoT also creates a vast web of devices and protocols that are vulnerable in various ways. However, cloud services are also used in IoT networks. This research concentrates only on cloud assets, although similar measurements can be used in future research in these mentioned areas.

5 RESULTS OF THE STUDY

We have now gone through vulnerabilities and threats that exist when services and data are transferred to could-service providers. Based on the various research on cloud threats, it was proved that even if services, computation, and data is transferred or migrated to the cloud service provider, it does not mean that it is more secure than earlier - there has to be a separate security plan for the identified cloud services.

In Chapter 3, we have also created seven hypotheses (Table 1) regarding why it is important to identify cloud assets for security purposes and what things should be kept in mind when using cloud services.

Table 1 Seven hypotheses on cloud security.

Identify all your assets	If you do not know your assets, it is impossible to protect yourself.
Acknowledge the dark web	By knowing your assets, you will get a better defence against ever-increasing criminal activities on the dark web.
Secure your IoT data	In the near future, there are an increasing amount of data and devices in the physical world, and these are all assets to be aware of.
Know your Cloud Service Provider	By knowing your assets, you know your cloud service provider assets and this way you control your own data.
Get familiar with your people	Human resources are also your assets. By knowing your cloud provider, you also need to know who works with your credentials.
Manage the whole supply chain	Knowing your assets means that you know, defend and control your whole supply chain.
Control the access	Someone you do not know already has your

to your data	assets.

As a result, a cloud asset identification strategy was established, and the right open-source tools were chosen for the task. After choosing an organisation for the strategy, several ways of determining its domains were explained. Domains of a registered company can be found on government registers and websites that provide that service, for example, the Cyber Exposure Index. After the domains are found, IP addresses that are linked with those domains are recognised. This part demands more work because of dynamic and passive IP addresses, sibling or shared domains and changing IP ranges. There are many ways to do this one address at a time, but to save time, there are websites that reduce time to the task at hand, for example, Robtex. Depending on the size of the company, there may be hundreds of IP addresses and domains linked to them and vice versa which makes it time consuming and careful work. In this research, the CIH platform was used to dig vulnerable IP addresses for ten listed companies' vulnerable domains.

After all IP addresses behind domains were recognised, other open-source tools were used to identify the 11 biggest cloud service providers. These tools are called IP2provider and Shodan CLI. All IP addresses were checked to verify whether they were cloud or not, which created a list of IP addresses and a cloud asset percentage with named cloud providers.

After the vigorous work, we have individual and unique information about every company separately: their cloud percentage, the service providers they use, and their IP addresses and domains. This information helps us to create a unique cloud asset identification strategy for every organisation, with detailed information on securing every IP address with vulnerabilities. The whole process of identifying cloud assets is presented in Table 2.

Table 2 Cloud asset identification strategy.

Choose an organisation	Choose a target for the cloud asset identification
Find its all domains	Take advantage on internet sites to find organisation's all domains
Recognize IP addresses behind domains	Use open-source tools to find out that IP addresses belong to found domains
Identify cloud assets	Another open-source tool set (presented in this thesis) to identify cloud addresses
Calculate cloud asset percentage	Rough calculation to understand the number of services that exists in the cloud
Create a cloud asset	Secure all your assets, especially cloud assets

strategy	and keep in mind 7 hypotheses

The same IP information tools with some modification can be used for other research, for example, identifying IoT assets, IIoT, SCADA or even IP reputation against phishing attacks, to mention a few. With these open-source tools available together with the cloud identification strategy process introduced in this thesis, any organisation will save hundreds of hours and thousands or euros in cybersecurity costs and consultant work.

The research question was to identify what means are needed for an organisation to use cloud services in a safe way. As an answer, we have created hypotheses which should be answered when making a safer cloud service environment for ourselves. After these hypotheses are tackled, thesis introduced cloud asset identification strategy with six steps which gives a roadmap to use all your identified cloud services in a safer way.

6 DISCUSSION AND CONCLUSIONS

This thesis concentrated on the importance of identifying cloud assets. Several previous studies were introduced containing threats and vulnerabilities that cloud services have. We also established seven hypotheses for why everyone should recognise which services they are using and how.

As the previous studies brought out, cloud services face several threats on their security. Virtualization together with many API frameworks increase the possibility for DDoS attacks because more attack surface is available for hackers. Data confidentiality and auditability in crisis situations is important to handle well and the strategy for that may differ between cloud service providers. If a cloud service goes out of business, data integrity, backups and access to data may become an issue. Technology risks include problems with hardware, resource sharing, portability and software problems. Data itself needs to be secure and available for the organisation. Clous Service Alliance has made a list of top threats facing the cloud security as we went through in Chapter 2 of previous studies. All these threats create a need for identifying our cloud assets so that we can make sure that necessary security measures are taken into consideration.

In short, if assets are not known by an organisation, it is impossible to protect them.

By knowing cloud assets (as well as other assets), a better defence against ever-increasing criminal activity is achieved. In the future, more data will leak throughout the supply chain, IoT devices and the physical world, not to mention people working in every aspect of it. Knowing all these assets is a huge, essential task. If it is not understood what kind of operating models your cloud service providers have or how you can control your own data there, it is impossible to be safe. Human resources are also assets. Knowing how people behave and where they work helps defend credentials. In the

end, an organisation may already have assets that are not known, or they are being used by someone that is not known. This is a reason for having both proactive and reactive asset strategies in place.

As a result, cloud asset identification was established. An organisation was chosen, its domains were recognised, and IP addresses connected to them. With this information and the correct tools, cloud assets were recognised. Based on the information on IP addresses and cloud service providers, a strategy can be created for any chosen company. Limitations would be to decide which information and cloud providers are enough to secure safety, for there is no tool to recognise everything, with updating the cloud or any other IP information. Multiple tools are available for future research tasks, some of which can be found in Appendix A. It is also important to recognise that these tools use IPv4 addresses, not IPv6.

The most important thing is to know what assets we have. If we do not know our assets or people who have access to them, there is no defence that could save us.

6.1 Limitations

As a limitation, all open-source tools need to be tested within time, with a much larger database of cloud service providers. With these tools, only the 11 biggest cloud providers were tested. New cloud service providers are coming to market all around the world, which means that there is no comprehensive tool set available anywhere that has all the cloud providers or even tools that are updated regularly. Even the biggest ones like Shodan are lacking IP information, even from the biggest cloud service providers. Such a tool should be created, updated regularly, and modified to identify other vulnerabilities too. The limitation would be to decide how many cloud providers are recognised, or from which area. Different global markets have different cloud providers, clients, services, and attack types (for example physical attacks).

One more limitation is the cloud service provider called CloudFlare. It is an example of a new cloud service that masks its IP addresses in such a way that it cannot recognise what services are used. If these kinds of services are increasing, identifying cloud assets will be a bigger task.

6.2 Future research direction

More research should be done in predicting attacks against cloud or IoT assets, especially from the dark web, and identifying more

information behind IP addresses, for example, IP reputation, to prevent phishing attacks.

Another topic for additional research would be to better understand the kind of attacks done with IP information, especially against cloud providers. The dark web as a new platform for attacks, and how to predict them proactively with natural language processing algorithms and machine learning tools, is another research area.

7 REFERENCES

Accenture (2020). Securing the supply chain. Understanding and mitigating the security risks of modern enterprise supply networks. *Accenture.* https://www.accenture.com/_acnmedia/PDF-134/Accenture-Securing-The-Supply-Chain.pdf

Aldrigde, J., & D.-H. D. (2015). Cryptomarkets: The darknet as an online drug market innovation. *University of Manchester, University of Montreal, Final report to NESTA*.

AlMenda, O.M., Alzahrani, S.M. (2021). Cloud and Edge Computing Security Challenges, Demands, Known Threats, and Vulnerabilities. *Academic Journal of Research and Scientific Publishing,* 2(21)*.*

Amichai-Hamburger, Y. (2009). Personality, individual differences and internet use. *Oxford handbook of internet psychology.* Oxford: Oxford University Press.

Benbasat, I. et. al (1987). The Case Research Strategy in Studies of Information Systems Case Research. *MIS quarterly*, 11(3), 369–386.

Bermbach, D. *et al*., On the Future of Cloud Engineering. *2021 IEEE International Conference on Cloud Engineering (IC2E)*, 264-275.

Dalins, J. et al. (2018). Criminal motivation on the dark web: A categorisation model for law enforcement. *Digital Investigation*, 24, 62–71.

Dingledine, R., Mathewson, N., & Syberson, P. (2004). Tor: The second-generation onion router. *13*.

Dolliver, D., & Kennedy, J. L. (2016). Characteristics of drug vendors on the Tor network: A cryptomarket comparison. *Victims*

Offenders, An International Journal of Evidence-Based Research, Policy, and Practice, 11(4).

Elbahrawy, A. E. (2020). Collective dynamics of dark web marketplaces. *Nature*. Sci Rep 10, 18827.

El Kafhali, S., El Mir, I. & Hanini, M. (2021). Security Threats, Defense Mechanisms, Challenges, and Future Directions in Cloud Computing. *Arch Computat Methods Eng.*

Faridi, F. et. al (2021). Cloud computing approaches health care, *Materials Today: Proceedings,* 2021.

FBI's Internet Crime Complaint Center (IC3). (2019). *2019 Internet Crime Rep* https://www.ic3.gov/Media/PDF/AnnualReport/2019_IC3Report.pdf

Feilky, M., Shahrestani, A and Ramadass, S., A Survey of Botnet and Botnet Detection. *2009 Third International Conference on Emerging Security Information, Systems and Technologies*, 2009, 268-273.

Finklea, K. (2017, March 10). Dark web. *Congressional Research Service*. https://sgp.fas.org/crs/misc/R44101.pdf

Guitton, C. (2013). A review of the available content on Tor hidden services: The case against further development. *Computers in Human Behavior*, 29(6), 2805–2815.

HackerOne. (2021). The 2021 hacker report: Understanding hacker motivations, development and outlook. *HackerOne*. https://www.hackerone.com/resources/reporting/the-2021-hacker-report?ungated=

IBM Business Blog. What is the Internet of Things (IoT)? *IBM*. https://www.ibm.com/blogs/internet-of-things/what-is-the-iot/

INFOSEC. Hacking communities in the deep web. *INFOSEC.* https://resources.infosecinstitute.com/topic/hacking-communities-in-the-deep-web/

Jardine, E. (2018). Privacy, censorship, data breaches and internet freedom: The drivers of support and opposition to Dark Web technologies. *New Media and Society*, 20(9), 2824–2843.

Kaur, S., Kaur, G. (2021). Threat and Vulnerability Analysis of Cloud Platform: A User Perspective. *8th International Conference on Computing for Sustainable Global Development (INDIACom),* 533-539.

Kwon, K. H. et al. (2017). Crisis and collective problem solving in dark web: An exploration of a black hat forum. *Black Hat Forum*. 1-5.

Li, X., Zhu, L., Chu, X., & Fu, H. (2020). Edge computing-enabled wireless sensor networks for multiple data collection tasks in smart agriculture. Journal of Sensors, 2020.

Madarie, R. (2017). Hackers motivations: Testing Schwartz's theory of motivational types of values in a sample of hackers. *International Journal of Cyber Criminology*, 11(1), 78–97.

Masson, K., & Bancroft, A. (2017). Nice people doing shady things: Drugs and the morality of exchange in the darknet cryptomarkets. *International Journal of Drug Policy*, 58, 78–84.

Marks, G. (2021). A casino gets hacked through a fish-tank thermometer. *Entrepreneur*. https://www.entrepreneur.com/article/368943

Merritt, D., & Mullins, B. (2011). Identifying Cyber Espionage: Towards a Synthesis Approach. *Journal of Network Forensics*, *Edith Cowan University Security Research Institute,* 3(1), 48-59.

Moor, L., & Anderson, J. R. (2018). A systematic literature review of the relationship between dark personality traits and antisocial online behaviours. *Personality and Individual Differences*, 144, 40–55.

Moyer, E. (2013). NSA disguised itself as Google to spy, say reports. *CNET.* https://www.cnet.com/news/nsa-disguised-itself-as-google-to-spy-say-reports/

NCTA. (2014). Infographic: The growth of the Internet of Things. *NCTA.* https://www.ncta.com/whats-new/infographic-the-growth-of-the-internet-of-things

Neshenko, N. et al. (2019). Demystifying IoT security: An exhaustive survey on IoT vulnerabilities and a first empirical look on internet-scale IoT exploitations. *IEEE Communications Surveys and Tutorials*, 21(3), 2702–2733.

Niemelä, M. (2016). Anatomy of cyberattack. *eBook.* 2016.

Niemelä, M. (2019). Leaking supply chain: Information available to hackers as a driving force for attacks. *Mikko Niemelä*. 2019.

Niemelä, M., & Koistinen, P. (2020). Smiling security. The cyber security manager’s road to success. *Mikko Niemelä.* Lioncrest: 2019.

Nithiasree, B.L., Prakash, R., Shenbaga Sundar, R. (2021). A Survey on Cloud Security Threats and Solution for Secure Data in Data Stages. *2021 International Journal of Computer Techniques (IJCT).* 8(2).

OSINTme. (2021). How to investigate a massive phishing campaign. *OSINTme.* https://www.osintme.com/index.php/2021/12/06/how-to-investigate-a-massive-phishing-campaign/

Perrons, R. K. (2015). How the Energy Sector Could Get it Wrong with Cloud Computing. *Energy Exploration & Exploitation*, *33*(2), 217–226.

Priya, R.P., Chouchan, J.T. (2018). Security Issues in Cloud Computing and Existing Solutions- a Survey. *International Journal of Engineering & Technology*.

Roethlisberger, R.J. (1977). The elusive phenomena. *Harvard Business School, Division of Research.* Boston: Harvard Business School.

Sapienza, A. et al. (2018). Early warnings of cyber threats in online discussions. 2017 *IEEE International Conference on Data Mining Workshops (ICDMW),* 667–674.

Sarkar, S. et al. (2017). Predicting enterprise cyber incidents using social network analysis on the darkweb hacker forums. *The Cyber Defense Review*, 87102. https://www.jstor.org/stable/26846122

Schatz, M. C., Langmead, B., & Salzburg, S. L. (2010). Cloud computing and the DNA data race. *Nature biotechnology*, 28, 691–693.

Senko, L. Do You Know Where Your Data Is? Securing Your Multicloud Environment. *Forbes.* https://www.forbes.com/sites/forbestechcouncil/2021/11/12/do-you-know-where-your-data-is-securing-your-multicloud-environment/

Shaikh, R., Sasikumar, (2012). Security issues in cloud computing: A survey. *International journal of computer applications,* 44(19).

Shajan, A. and Rangaswamy, S., 2021. Survey of Security Threats and Countermeasures in Cloud Computing. *United International Journal for Research & Technology (UIJRT),* 2(7), 201-207.

Sinanović, H., Mrdovic, S., Analysis of Mirai malicious software. *2017 25th International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, 1-5.

Sultan, N. Cloud computing for education: A new dawn?, *International Journal of Information Management,* 30(2), 109-116.

Tahirkheli A.I. et.al. (2021). A Survey on Modern Cloud Computing Security over Smart City Networks: Threats, Vulnerabilities, Consequences, Countermeasures, and Challenges. *Electronics*. 10(15). 1811.

Tamber Ebot, A. C. (2021). Social engineering attacks in information systems. *Zicklin School of Business*, New York: Baruch College.

Traficom. Criteria to assess the information security of cloud services (PiTuKri). *Traficom.* Finnish transport and communications agency, National Cyber Security Centre Finland.

United Nations Office of Drugs and Crime. (2020). Darknet cybercrime threats to Southeast Asia. *United Nations*. https://www.unodc.org/documents/southeastasiaandpacific//Publications/2021/Darknet_Cybercrime_Threats_to_Southeast_Asia_report.pdf

Van Hout, M. C., & Bingham, T. (2013). Responsible vendors, intelligent consumers: Silk Road, the online revolution in drug trading. *International Journal of Drug Policy*, 25(2), 183–189.

Villalva, D. A. B. et al. (2018). Under and over the surface: A comparison of the use of leaked account credentials in the dark and surface web. *Crime Sci*, 7, 17.

Vipinraj, N. Top IoT trends and tech in 2021. *Vipinraj Nair.* https://vipinrajnair.medium.com/top-iot-trends-and-tech-in-2021-f52766cf2414

Wall, D. S. (2011). Crime and the Internet. *Routledge*. 2011.

Warkenting, M. et al. (2012). The role of individual characteristics on insider abuse intentions. *AMCIS 2012 Proceedings*, 28. https://aisel.aisnet.org/amcis2012/proceedings/ISSecurity

8 APPENDIX 1

A list of tools for analyzing internet traffic.

Tool	Description & method	Feasibility	Source
which-cloud	Given an ip address, return which cloud provider it belongs to (EC2, GCE, etc)	Install: npm i which-cloud -g. usage: which-cloud 52.4.0.0	https://github.com/bcoe/which-cloud
Cloud_ip_ranges	Most cloud providers publish up to date lists of their IP address ranges. This tool identifies if an IP belongs to a provider's ranges by fetching and parsing the latest lists.	Install: download the zip package, then from the root: pip install -r requirements.txt. usage, e.g. Python cloud_ip_ranges 52.4.0.0	https://github.com/nccgroup/cloud_ip_ranges
ip2provider	Check which cloud provider is hosting a particular IP address. Some providers will have service and region listed	Install: pip3 install -r requirements.txt . usage e.g. ./ip2provider.py 52.4.0.0	https://github.com/oldrho/ip2provider
Shodan	Expose all open IP addresses and ports in the internet, devices services etc. It has a CLI version.	Install a CLI version: pip install -U --user shodan, initialise with your API key: shodan init YOUR_API_KEY, usage: e.g. download the most recent 200 results for aws: shodan download --limit 200 myresults.json.gz aws. OR: Show a comma-separated list of aws IPs and ports: shodan	shodan.io

		search --fields ip_str,port --separator , aws	
server-ip-addresses	Daily updated list of IP addresses / CIDR blocks used by data centres, cloud service providers, servers, etc.	Essentially the IP addresses where an average web user should not be accessing from. Useful for detecting or limiting traffic from servers.	https://github.com/jhassine/server-ip-addresses
Nikto	Good additional tool for cloud vuln scanning. Web server vuln scanner, performs tests against Web servers for 6700 potentially dangerous files/programs, checks for outdated versions of over 1250 servers and version-specific problems on over 270 servers, checks for server configuration items, e.g. presence of multiple index files and http server options, will attempt to identify installed web servers and software, frequently updated. especially good with server misconfiguration	Install: git clone https://github.com/sullo/nikto. usage: cd nikto/program/ , ./nikto.pl -h http://www.jyu.fi	https://securitytrails.com/blog/nikto-website-vulnerability-scanner
ScoutSuite	Good additional tool for checkin your cloud account's configurations. Multi-cloud security auditing tool for AWS Google Clouds and Azure environmants (Python)	More advanced to install, easy to use, e.g. scout AWS --profile basc -f. you need cloud accounts.	https://github.com/nccgroup/ScoutSuite
Nmap	Instead of pinging one device, you can ping the entire network. nmap sends packets and reads responses to discover hosts and services across the network. ping scan, port scanning, nmap stealth mode, analysing with wireshark, detect	Easy to use and install, although it requires knowledge of terminal use	https://nmap.org/

	operating systems, aggressive mode, decoy usage, etc. Examples: nmap 192.168.1.1. with many different options, is the host up, which ports are used, TCP or UDP, open or filtered, what service is running (domain, http, netbios, microsoft-ds etc.)		
OpenVAS	Software framework of several services for vuln mgmt, built to be an all-in-one scanner, runs from a security feed of over 50 000 vuln tests, updated daily.	For more advanced users, even installation is difficult for a basic user, you need to use c compilers, cmake and many other software or libraries of certain versions. Good for experienced users who want to perform target scans or pen-testing; installing and using it has a significant learning curve and needs an experienced admin.	https://openvas.org/
AWS Public IPs	Fetch all public ip addresses tied to your AWS account, works with IPv4 / IPv6 classic / VPC networking and across all AWS services (ruby)	Seems to have quite a lot of options	https://github.com/arkadiyt/aws_public_ips
PMapper	Advanced and automated AWS IAM evaluation (python)	Command-line tool, e.g. # Create a graph for the account, accessed through AWS CLI profile "skywalker" pmapper --profile skywalker graph create # [... graph-creation output goes here ...]	https://github.com/nccgroup/PMapper
nccgroup AWS-Inventory	Make an inventory of all your resources across regions (python)	Easy. Running with defaults: python aws_inventory.py	https://github.com/nccgroup/aws-inventory
Ettercap	Packet capture utility and an attack facility can inject traffic into a stream, command-line utility that has an	Maybe for Advanced users, you need to install dependencies unless you have kali linux where it seems to	https://www.ettercap-project.org/

	extensive library of commands, attacks are based on the man-in-the-middle strategy, methods rely on the utility being run from within a network. Password capture, DNS spoofing, Denial of Service	be installed. It is like a metasploit console, a customised Command Promt / Terminal Window, so you need at least a basic understanding of working with bash. You need to set up your computer to allow packet flow to your computer, e.g. sysctel -w net.ipv4.ip_forward=1 if you are acting between router and the client. Example: you make an arp poisoning attack by giving two targets and poisoning them. After that, you have a man in the middle attack, which you can monitor with wireshark or tcpdump.	
Sqlmap	Command-line utility offers just one command but has hundreds of options which changes the function that gets run, documents databases and launches attacks. Password cracking, injection attacks	sqlmap basically automates the reconnaissance and exploitation of database. Its purpose is to print out or enumerate existing databases. you can use this together with burp suite. For example, you can intercept the request of a web client to the server for log in. Put on intercept from burp suite, then use local host proxy, then you capture that request and use sqlmap to enumerate the database. usage: sqlmap -r /whatever/request.txt, then you choose the injection point, and as a result you get server operating system, application technology, back-end DBMS, etc.	https://sqlmap.org/
OWASP ZAP	Integrated tool for finding vulns in Web apps, a fork of the Paros Proxy Tool,	Easy. from the graphical ui, just click quick start and automated scan, then give an address,	https://www.zaproxy.org/

	provides automated scanners, a set of tools for finding security vulns manually, supported by many orgs incl. OWASP, Microsoft and Google. Like burp suite, but completely free. Comes as a default in kali linux	and press attack. All other options can be found from drop-down menus.	
Clair	Specialised container vuln analysis service, provides a list of vulns that may threaten a container, analyses each container once and does not execute container to perform its examination, extracts all required data to detect known vulns and caches layer data for examination against vulns discovered in the future. works with these or supports these: ubuntu, debian, RHEL, Suse, Oracle, Alpine, AWS Linux, VMWare Photon, Python	This is more difficult and advanced, seems to work with manifests or something. The simplest way to submit a manifest to your running Clair is to utilise clairctl. This is a CLI tool capable of grabbing image manifests from public repositories and submitting them for analysis. Clair's analysis has three parts: indexing (submit manifest to Clair, it fetches layers, scans their contents, and returns IndexReport), matching (takes IndexReport and correlates vulns affecting the manifest) and notifications (when a new vuln is discovered, the notifier determines whether this vuln affects any indexed Manifests).	https://github.com/coreos/clair
Arkime	Large scale IPv4 packer capturing (PCAP), indexing and database system, PCAP browsing, searching, and exporting. Doesnt replace IDS engines but works alongside them to store and index all the network traffic in standard PCAP format. Has three components: capture (threaded C app monitoring network	Once Arkime is running, point your browser to http://localhost:8005 to access the web interface. Click on the Owl to reach the Arkime Help page.	https://arkime.com/

	traffic, writes PCAP formatted files to disk, parses the captured packets, sends metadata (SPI) to elasticsearch), viewer (node.js app runs per capture machine, handles the web interface and transfer pcap files) and elasticsearch (search database technology)		
Powerfuzzer	Automated and customisable web fuzzer (http protocol-based app fuzzer), capable of identifying these problems: cross site scripting XSS, injections SQL LDAP code commands CRLS XPATH, http500 statuses	More advanced, need to have understanding about fuzzing. The source code is commented on so you could get information from there; otherwise, there is not much documentary	https://www.powerfuzzer.com/
arachni-scanner	Modular high-performance Ruby framework aimed to help pen-testers and admins evaluate the security of modern web apps, multi-platform, supporting all major OS's, command line scanner utility, high-performance global scanners, Ruby library allowing for scripted audits, multi-user multi-scan web collaboration platform, based on REST API, supports highly complicated web apps	There are no dependencies like databases, system services, libraries, or any configuration overhead. Simply download and extract one of our packages to a supported OS and run a script, a scan, and fire-up the web interface or convert the machine to a Grid node – all with a single command.	https://www.arachni-scanner.com/
Deepfence ThreatMapper	Monitor and secure running apps. Kubernetes, Docker, bare metal and VM-based platforms, AWS Fargate. Discover running workloads, vulns and rank vulns by risk-of-exploit.	More advanced, needs to install sensors on kubernetes, docker hosts etc	https://github.com/deepfence/ThreatMapper
Grabber	Web apps scanner,	More advanced. you	https://rgaucher.info/

	detects vulns in your website, simple, not fast but portable, adaptable, designed to scan small websites like forums	need to install and configurate, then the actual using is just a python script, e.g. python grabber.py --spider 1 --sql --xss --url http://127.0.0.1/bank	beta/grabber/
Grendel-Scan	Web app sec testing tool, automated testing module for detecting common web app vulns, and features geared at aiding manual pen tests	More advanced, unorganised tips of use in github	https://sourceforge.net/p/grendel/code/ci/c59780bfd41bdf34cc13b27bc3ce694fd3cb7456/tree/
Nuclei	Write your own custom security checks	Quite easy, works from terminal. nuclei -h gives options	https://nuclei.projectdiscovery.io/
sec-helpers	Collection of dynamic security related helpers (DAST), bundle of useful tests and validators to ensure the security of a given domain	Copy and change the following to run all tests: import sec_helpers domain: str = 'vwt-digital.github.io' # {domain}.{tld} sec_helpers.CorsPolicy(domain=domain) sec_helpers.HighTls(domain=domain, slide=False) # Slide is False by default sec_helpers.Hsts(domain=domain, age=10368000) # Age is 10368000 by default sec_helpers.NoHttp(domain=domain) sec_helpers.NoSsl(domain=domain)	https://pypi.org/project/sec-helpers/
Wapiti	Web app vuln scanner, performs black-box scans, meaning it does not study the source code, but by crawling the webpages looking for scripts and forms it can inject data. Once it gets the list of URLS, forms and their inputs, it acts like a fuzzier injecting payload to see if a script is vulnerable.	Quite easy, command-line tool. usage: wapiti [-h] [-u URL] [--scope {page,folder,domain,url,punk}] [-m MODULES_LIST] [--list-modules] [--update] [-l LEVEL] [-p PROXY_URL] [--tor] [-a CREDENTIALS] [--auth-type {basic,digest,ntlm,post}] [-c COOKIE_FILE] [--drop-set-cookie] [--skip-crawl] [--resume-crawl]	https://wapiti.sourceforge.io/

		[--flush-attacks] [--flush-session] [--store-session PATH] [--store-config PATH] [-s URL] [-x URL] [-r PARAMETER] [--skip PARAMETER] [-d DEPTH] [--max-links-per-page MAX] [--max-files-per-dir MAX] [--max-scan-time SECONDS] [--max-attack-time SECONDS] [--max-parameters MAX] [-S FORCE] [--tasks tasks] [-t SECONDS] [-H HEADER] [-A AGENT] [--verify-ssl {0,1}] [--color] [-v LEVEL] [-f FORMAT] [-o OUTPUT_PATH] [--external-endpoint EXTERNAL_ENDPOINT_URL] [--internal-endpoint INTERNAL_ENDPOINT_URL] [--endpoint ENDPOINT_URL] [--no-bugreport] [--version]	
w3af	Web app attack and audit framework, using python, easy to use and extend	The framework has three main plugin types: crawl (find new URLs forms and other injection points e.g. web spider), audit (take the injection points and send specially crafted data to all in order to identify vulns) and attack (exploit vulns found by audit plugins, usually return a shell on the remote server or a dump of remote sql tables)	https://w3af.org/
Prowler	CIS benchmarks and additional checks for security best practises in AWS (bash and python components	A bit more advanced need to have understanding about clouds and dockers. command line would be e.g. ./prowler -p custom-profile -r us-east-1 to create AWS-CLI profile	https://github.com/toniblyx/prowler

		and region	
CloudSploit Scans	AWS security scanning checks (NodeJS)	More advanced, different options to different clouds, e.g. AWS, Azure, Google Cloud, Oracle Cloud etc	https://github.com/aquasecurity/cloudsploit
CloudMapper	Helps you analyse your AWS environments (python)	Python script. run with demo data: # Generate the data for the network map Python cloudmapper.py prepare --config config.json.demo --account demo # Generate a report Python cloudmapper.py report --config config.json.demo --account demo Python cloudmapper.py webserver	https://github.com/duo-labs/cloudmapper
CloutTracker	Find over-privileged IAM users and roles by comparing CloudTrail logs with IAM policies (python)	Quite easy, although some setup needs to be done before using. usage: cloudtracker --account demo --list users	https://github.com/duo-labs/cloudtracker
AWS security benchmarks	Scripts and templates guidance related to the AWS CIS foundation framework (python)	No documentation	https://github.com/amazon-archives/aws-security-benchmark
Resource Counter	Counts number of resources in categories across regions	Easy to run. Python count_resources.py with options	https://github.com/disruptops/resource-counter
ICE	Insights from a usage and cost perspective with high detail dashboards	Looks quite complicated	https://github.com/Teevity/ice
SkyArk	Advanced discovery and security assessment for the most priviledged entities in the testes AWS	A bit more complicated, I don’t know if this is a PowerShell tool	https://github.com/cyberark/SkyArk
Trailblazer AWS	Determine what AWS API calls are logged by CloudTrail and what they are logged as, can be used as an attack simulation	Some setup needs to be done first, but the actual use is quite easy: trailblazer --help	https://github.com/willbengtson/trailblazer-aws

	framework		
Lunar	Security auditing tool based on several security frameworks, does some AWS checks	Easy: ./lunar.sh [OPTIONS...]	https://github.com/lateralblast/lunar
Cloud-reports	Scans your AWS cloud resources and generates reports	Run and generate html report: npm run scan -- --profile Your-AWS-profile -f html	https://github.com/tensult/cloud-reports
Pacbot	Platform for continuous compliance monitoring, reporting and sec automation for the cloud	Quite easy to use but you need understanding of AWS	https://github.com/tmobile/pacbot
cs-suite	Integrates tools like Scout2 and Prowler together	Normal Python script. usage: cs.py [-h] -env {aws,gcp,azure,digitalocean} -aip AUDIT_IP -u USER_NAME -pem PEM_FILE [-p] [-pId PROJECT_ID] [-az_u AZURE_USER] [-az_p AZURE_PASS] [-o OUTPUT] [-w] [-n NUMBER]	https://github.com/SecurityFTW/cs-suite
AWS-key-disabler	Small lambda script that disables access keys older than a given number of days	More advanced, need lambda understanding	https://github.com/te-papa/aws-key-disabler
Antiope	AWS inventory and compliance framework	Looked quite difficult for me at least	https://github.com/turnerlabs/antiope
Cloud Reports	Scans your AWS cloud resources and generates reports and includes security best practises	Quite easy, although needs additional setup, e.g. npm run scan -- --profile Your-AWS-profile	https://github.com/tensult/cloud-reports
Terraform AWS Secure Baseline	Set up your AWS account with the secure	More advanced, needs some programming	https://github.com/nozaq/terraform-aws-secure-baseline
Cartography	Python tool, consolidates infra-assets and the relationships between them in an intuitive graph view powered by Neo4j database	Looked quite advanced instructions	https://github.com/lyft/cartography
TrailScrap	Command-line tool to	With a command trail	https://github.com/

er	get valuable information out of AWS CloudTrail	scraper and options for many ways to use this, easy to use but a lot of options	flosell/trailscraper
LambdaGuard	An AWS lambda auditing tool designed to create asset visibility and provide actionable results	Quite easy. lambdaguard --help	https://github.com/Skyscanner/LambdaGuard
Komiser	Environment inspector, analyse and manage cloud cost usage security and governance in one place	More advanced, many steps to use	https://github.com/mlabouardy/komiser
Perimeterator	AWS perimeter monitoring, periodically scan internet facing AWS resources to detect misconfigured services	Couple of tools in the same package, a bit more advanced	https://github.com/darkarnium/perimeterator
PolicySentry	IAM lead privileged policy generator auditor and analysis database	Many steps in the using, more advanced	https://github.com/salesforce/policy_sentry
Zeus	AWS auditing and hardening tool	Instructions a bit messy, a bit more advanced	https://github.com/DenizParlak/Zeus
janiko71 AWS-inventory	Python script for AWS resources inventory	No documentation or clear instructions	https://github.com/janiko71/aws-inventory
awspx	Graph-based tool for visualising effective access and resource relationships in AWS environments	Simple to use: awspx ingest	https://github.com/fsecurelabs/awspx
clinv	DevSecOps command-line asset inventory tool	No simple instructions ready, but looks quite doable	https://github.com/lyz-code/clinv
AWS-gate	Enhanced AWS SSM session manager CLI client	Need telecommunication understanding	https://github.com/xen0l/aws-gate
Detecting Credential Compromise	Detecting of your compromised credentials in AWS	Quite easy to use, not to easy to understand. detect [OPTIONS]	https://github.com/Netflix-Skunkworks/aws-credential-compromise-detection
AWS-Security-Toolbox	AWS security toolbox (docker image) for security assessments	This looked more advanced	https://github.com/z0ph/aws-security-toolbox

(AST)

IAM-LINT	GitHub action for linting AWS IAM policy documents for correctness and possible security issues	I didn’t understand this right away so more advanced	https://github.com/xen0l/iam-lint
AWS-security-viz	Tool to visualise AWS security groups	Didn’t look too easy to use. Generate the graph directly using AWS keys aws_security_viz -a your_aws_key -s your_aws_secret_key -f viz.svg --color=true	https://github.com/anaynayak/aws-security-viz
AirIAM	Least privilege AWS IAM using Terraform	Easy to use, harder to understand. usage: airiam find_unused [-h] [-p PROFILE] [-l LAST_USED_THRESHOLD] [--no-cache] [-o {cli}]	https://github.com/bridgecrewio/AirIAM
Cloussplaining	AWS IAM sec assessment tool, identifies violations of least privilege and generates a risk-prioritised HTML report	Like many others, setup is more difficult than use. e.g. cloudsplaining scan-policy-file --input-file examples/policies/explicit-actions.json	https://github.com/salesforce/cloudsplaining
IAM-policy-generator	A simple library to generate IAM policy statements with no need to remember all the actions APIs	Need programming skills	https://github.com/aletheia/iam-policy-generator
SkyWrapper	Helps to discover suspicious creation forms and uses of temporary tokens in AWS	Not so easy, programming understanding needed	https://github.com/cyberark/SkyWrapper
AWS-recon	Multi-threaded AWS inventory collection tool	Quite easy to use, e.g. $ AWS-vault exec profile -- AWS_recon	https://github.com/darkbitio/aws-recon
IAM-policies-cli	CLI tool for building simple to complex IAM policies	Easier to use than understand. iam-pol -t template.yaml -f yaml	https://github.com/mhlabs/iam-policies-cli
Aaia	AWS identity and access management visualiser and anomaly finder	Instructions looked more advanced	https://github.com/rams3sh/Aaia
IAM-floyd	IAM policy statement	I didn’t find any	https://github.com/

	generator with fluent interface - available for Node.js, Python, .Net and Java	instructions easily	udondan/iam-floyd
rpCheckup	AWS resource policy security checkup tool that identifies public external account access, intra-org account access and private resources	Easy to use: ./rpCheckup . but the results need understanding	https://github.com/goldfiglabs/rpCheckup
S3 Exif Cleaner	Remove EXIF data from all objects in an S3 bucket	Quite easy: s3_cleanse.py [-h] [-b BUCKET] [-p PREFIX]	https://github.com/seisvelas/S3-Exif-Cleaner
Steampipe	Use SQL to instantly query your cloud services (AWS, Azure, GCP and more), open-source CLI, no DB required (SQL)	Quite easy to use, need understanding on SQL., e.g. steampipe query	https://github.com/turbot/steampipe
WeirdAAL	AWS attack library	I did not find easy instructions	https://github.com/carnal0wnage/weirdAAL
Pacu	AWS penetration testing toolkit	Need understanding, e.g. about dockers, AWS	https://github.com/RhinoSecurityLabs/pacu
Cred Scanner	Simple file-based scanner to look for potential AWS access and secret keys in files	Normal Python script: Python cred_scanner.py. results require understanding	https://github.com/disruptops/cred_scanner
AWS PWN	Collection of AWS penetration testing junk	Many different options, more advanced	https://github.com/dagrz/aws_pwn
Cloudfrunt	Tool for identifying misconfigured CloudFront domains	Python script: cloudfrunt.py [-h] [-l TARGET_FILE] [-d DOMAINS] [-o ORIGIN] [-i ORIGIN_ID] [-s] [-N]	https://github.com/MindPointGroup/cloudfrunt
Cloudjack	Route53/CloudFront vuin assessment utility	Easy command results more difficult to understand. Python cloudjack.py -o json -p default	https://github.com/prevade/cloudjack
Nimbostratus	Fingerprinting and exploiting Amazon cloud infra	No available instructions	https://github.com/andresriancho/nimbostratus
GitLeaks	Audit git repos for secrets	Easy to use, but many options, e.g. gitleaks	https://github.com/zricethezav/gitleaks

		[OPTIONS]	
TruffleHog	Searches through git repositories for higg entropy strings and secrets digging deep into commit history	Command line basic tool: usage: trufflehog [-h] [--json] [--regex] [--rules RULES] [--allow ALLOW] [--entropy DO_ENTROPY] [--since_commit SINCE_COMMIT] [--max_depth MAX_DEPTH] git_url	https://github.com/trufflesecurity/truffleHog
Dumpster Diver	Tool to search secrets in various filetypes like keys (e.g. AWS access key, Azure share key, SSH keys) or passwords	Normal command line tool. usage: DumpsterDiver.py [-h] -p LOCAL_PATH [-r] [-a] [-s] [-o OUTFILE] [--min-key MIN_KEY] [--max-key MAX_KEY] [--entropy ENTROPY] [--min-pass MIN_PASS] [--max-pass MAX_PASS] [--pass-complex {1,2,3,4,5,6,7,8,9}] [--exclude-files EXCLUDE_FILES [EXCLUDE_FILES ...]] [--bad-expressions BAD_EXPRESSIONS [BAD_EXPRESSIONS ...]]	https://github.com/securing/DumpsterDiver
Mad-King	Proof of concept zappa based AWS persistence and attack platform	More advanced, e.g. Clone the repository git clone git@github.com:ThreatResponse/mad-king.git Set up a virtualenv in the directory. virtualenv. Activate source bin/activate. Install the requirements pip install -r requirements.txt Set up a boto profile in your ~/.aws folder for the account you are "hacking" zappa deploy production	https://github.com/ThreatResponse/mad-king
Cloud-Nuke	Tool for cleaning up your cloud accounts by nuking (deleting) all resources within it	Easy use, more options available, e.g. cloud-nuke aws	https://github.com/gruntwork-io/cloud-nuke
MozDef-The Mozilla	Seeks to automate the security incident handling process and	Many options, big instructions, more advanced	https://github.com/mozilla/MozDef

Defence Platform	facilitate the real-time activities of incident handlers		
Lambda-Proxy	Bridge between SQLMap and AWS lambda which lets you use SQLMap to natively test AWS lambda functions for SQL injection vulns	Easy to use, harder to get it. python3 main.py	https://github.com/puresec/lambda-proxy
CloudCopy	Cloud version of the Shadow Copy Attack against domain controllers running in AWS using only the EX2:CreateSnapshot permission	More advanced	https://github.com/Static-Flow/CloudCopy
enumerate-IAM	Enumerate the permissions associated with AWS credential set	More advanced	https://github.com/andresriancho/enumerate-iam
Barq	A post-exploitation framework that allows you to perform attacks on a running AWS infra	More advanced	https://github.com/Voulnet/barq
CCAT	Cloud container attack tool for testing security of conainter environments	More advanced, related to cloud service	https://github.com/RhinoSecurityLabs/ccat
Dufflebag	Search exposed EBS volumes for secrets	More advanced	https://github.com/bishopfox/dufflebag
attack_range	Tool that allows you to create vulnerable instrumented local of cloud environments to simulate attacks against and collect the data into Splunk	Easy to use but many steps, e.g. python attack_range.py configure	https://github.com/splunk/attack_range
whispers	Identify hardcoded secrets and dangerous behaviours	Quite easy, though many options. whispers --help	https://github.com/Skyscanner/whispers
Redboto	Red Team AWS Scripts	A lot of tools, advanced	https://github.com/elitest/Redboto
CloudBrute	A tool to find company (target) infra, files, and apps on the top cloud providers	Quite easy. usage: CloudBrute [-h|--help] -d|--domain "<value>" -k|--keyword "<value>" -w|--wordlist "<value>" [-c|--cloud "<value>"] [-	https://github.com/0xsha/cloudbrute

		t|--threads <integer>] [-T|--timeout <integer>] [-p|--proxy "<value>"] [-a|--randomagent "<value>"] [-D|--debug] [-q|--quite] [-m|--mode "<value>"] [-o|--output "<value>"] [-C|--configFolder "<value>"]	
AWS IR	AWS specific incident response and forensics tool	Command line tool this one too. usage: aws_ir [-h] [--version] [--verbose] [--profile PROFILE] [--case-number CASE_NUMBER] [--examiner-cidr-range EXAMINER_CIDR_RANGE] [--bucket-name BUCKET_NAME] [--dry-run] {instance-compromise,key-compromise} ...	https://github.com/ThreatResponse/aws_ir
CloudQuery	Exposes cloud configuration and metadata as sql tables, analysis and monitoring for compliance and security	This is more advanced tool, requires understanding of cloud setups	https://github.com/cloudquery/cloudquery/
AWS_Responder	AWS digital forensic and incident DFIR response, python scripts	AWS understanding needed. usage: aws_respond.py [-h] [--module MODULE] [--listmodules] [--moduledetails] [--dryrun DRYRUN] [--instanceids INSTANCEIDS [INSTANCEIDS ...]] [--sgids SGIDS [SGIDS ...]] [--vpcids VPCIDS [VPCIDS ...]] [--usernames USERNAMES [USERNAMES ...]] [--accesskeyids ACCESSKEYIDS [ACCESSKEYIDS ...]] [--values VALUES [VALUES ...]]	https://github.com/prolsen/aws_responder

Cloud-forensics-utils	Python library to carry out DFIR analysis on the google cloud	More advanced, many tools and options	https://github.com/google/cloud-forensics-utils
Dagda	Perform static analysis of known vulns, malware, viruses, trojans and other malicious threats in docker images of containers	Easy to use, but many options, e.g. python3 dagda.py check --docker_image DOCKER_IMAGENAME	https://pentestit.com/dagda-docker-security-suite/
dirb	Web content scanner that looks for hidden directories and files on the target website	No instructions whatsoever easily at least	http://dirb.sourceforge.net/
WPScan	WordPress vuln scanner	More advanced	https://github.com/wpscanteam/wpscan
IPInfo	Information about IP addresses	Search web site	https://ipinfo.io/
IPv6 port scanner	Port scanner web site	Just insert an ip address	http://www.ipv6scanner.com/cgi-bin/main.py
TCP Open port Scanner	Open port scanner web site	Scan ports with ip address web site	https://gf.dev/port-scanner
Mirage	Security audits and pentesting	More advanced	https://homepages.laas.fr/rcayre/mirage-documentation/overview.html
Kismet	Kismet is a wireless network and device detector, sniffer, wardriving tool, and WIDS (wireless intrusion detection) framework	More advanced	https://www.kismetwireless.net/
Spiderfoot	OSINT finder	More advanced	https://www.spiderfoot.net/
Scout2	Security auditing tool for AWS environments	Easy to use but many options. Scout2	https://github.com/nccgroup/Scout2
dsniff	Collection of tools for network auditing and penetration	More advanced, many tools in one package	https://github.com/tecknicaltom/dsniff