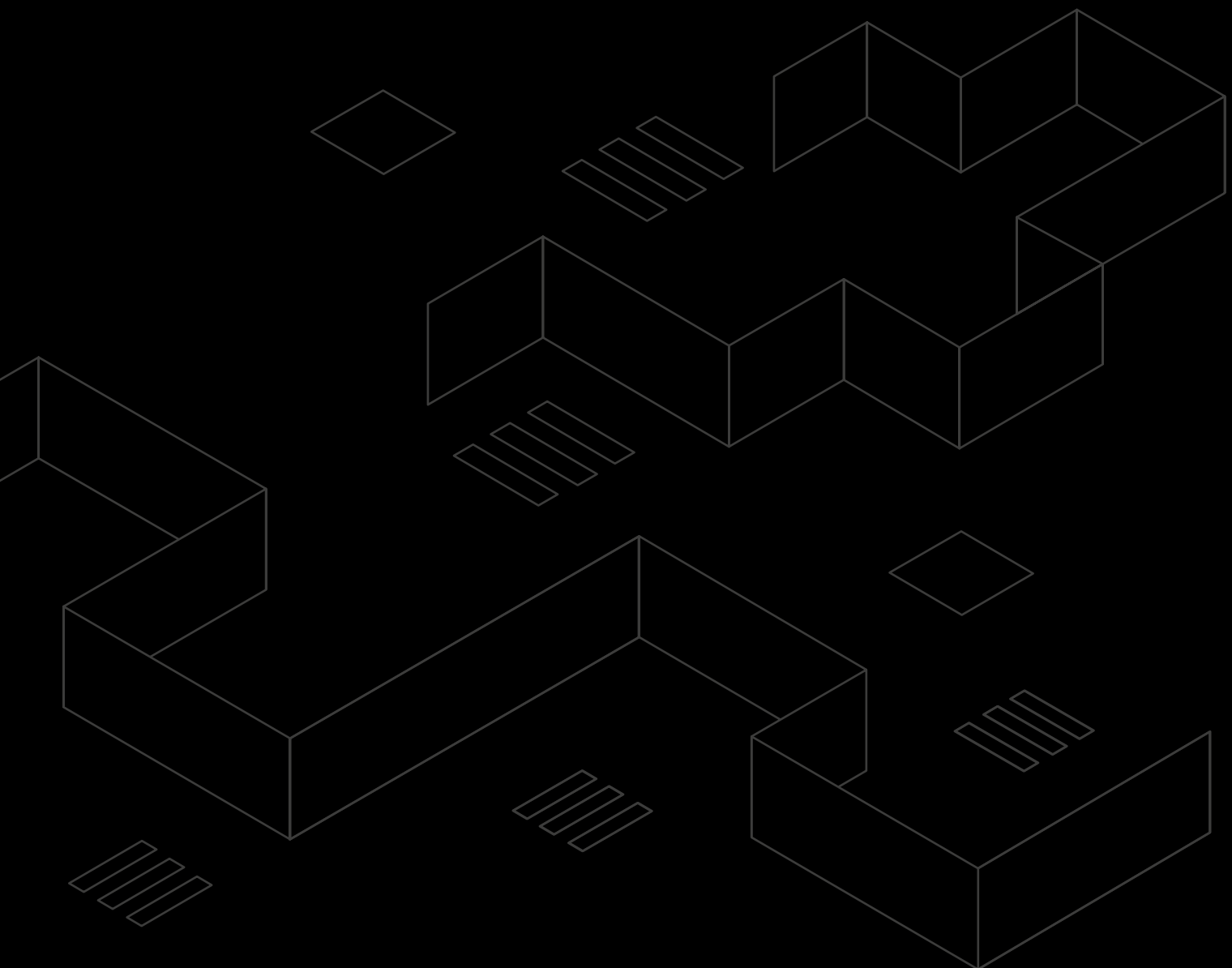
mnemonic

Securing your business.

2021

SECURITY REPORT



SECURITY REPORT 2021

In 1983, Ken Thompson received the ACM Turing Award. In his acceptance speech "Reflections on Trusting Trust", he outlined what has later been called "the most subversive of all hacks": He had modified the UNIX login program to accept either the intended encrypted password for the user, or a hard-coded password of his own choosing. Thompson had further modified the C compiler (cc) to detect whether it was compiling the login program and then to insert the backdoor automatically. He even made the compiler detect if it was compiling a new version of itself, and then automatically add both the code for modifying the login program, and the self-perpetuating code to the compiler. Thompson thus created a fully self-sustaining trojan, invisible to any code-reviewers.

The recent Solarwinds compromise exhibits very similar traits: The attackers, having gained access to the build system at Solarwinds, were able to inject the malicious code into the finished product without leaving any traces in the source code. They achieved this by looking for process instances of MsBuild.exe (the build tool in Visual Studio), and swapping out a particular source file for their own "bugged" version on the fly at build time. They took great care to suppress any errors from the build process to avoid being detected by the programmers. (They also used a raft of other fairly advanced techniques, too many to describe here.) The type of scenario that Thompson described decades ago has thus fully become reality.

We can refer to this type of attack as "supply chain", meaning that the intended victims were customers of Solarwinds, rather than the company itself. But even though we have a name for it, it is still unprecedented in a number of ways. Particularly the degree of stealth and the scope of the breach they achieved is notable. Furthermore, the attackers went after another foundational aspect of digital security, namely *identity*. For the victims where trojanised Solarwinds installations were leveraged as an intrusion vector, the attackers created "golden SAML" tokens for lateral movement and persistence, giving them wide access to tenants in Azure (and more). All of this reflects back to Thompson's "trusting trust": If we cannot trust software even from large, reputable companies, and we cannot trust digitally signed access tokens, then we are in an even more serious predicament than previously acknowledged.

The good news is of course that all the capabilities we as a company have built these last 20 years are directly applicable. The comprehensive nature of our approach is helpful: Our systems and procedures for security monitoring, threat intelligence, vulnerability scanning and threat hunting have all been deployed to good effect since the breach became known. Our various consulting departments have been able to analyse, advise and strengthen the defences of customers. We will surely keep developing all of these capabilities as the nature of the threats evolve.

That said: Solarwinds still presents us with serious questions. We still do not know the full extent of this operation. The information from victims (and vendors) is curiously incomplete. We do not know the intention behind it: Was it "merely" espionage or a dress rehearsal for something far more sinister?

In conclusion: Solarwinds have certainly made our work more interesting. And it has made our critical role in society even clearer.

I hope you will enjoy this edition of our security report. Though I am sure it still contains "bugs", I can certainly promise that none of them will be even slightly subversive!

TØNNES INGEBRIGTSEN
CEO, mnemonic

TABLE OF CONTENTS SECURITY REPORT 2021

ARTICLES

ARTICLES

INTERVIEWS I STATISTICS


2021
SECURITY
PREDICTIONS

Morten Weea
Senior Threat Intelligence Researcher

The time of year has come to close the chapter on the previous year and gaze into the horizon which is the coming year.

The philosopher George Santayana once said: "To know your future you must know your past", and 2020 has proven to be our black swan in many ways. To know our past, we will therefore first take a brief look at what we predicted for 2020, and how our predictions turned out to fit with reality.

A glance back at 2020

2020 was an unusual year. Not necessarily because we didn't accurately predict 2020, but because it was the year of many disruptions and involuntary drivers of change and digitalisation.

In last year's report, we spoke about the further professionalisation of cyber criminals. Compared to what was to come in 2020, the year started off fairly quiet – perhaps an ominous sign of the global turmoil that would strike in March. COVID-19 hit like a runaway freight train. Countries went into lockdown and companies graciously took their final bow one after the other. The world collectively started to shift its focus towards finding a solution to the ongoing pandemic, and there were even signs of a fragile truce from some cyber adversaries. A truce like this indicates some sort of coordination and pro-fessionalisation. Some of the more prominent adversaries even published statements that they would stay away from selected businesses for some time.

As time went on and the world started to live with the pandemic being the new normal, the situation became politicised and the truce slowly but steadily weathered.

The pandemic has also been an unmatched driver for digitalisation. Not only were office workers suddenly mandated to work from home, but country-wide lockdowns forced companies to adapt and ensure their services could be delivered digitally wherever possible. Naturally, there were some temporary hacks to allow employees to work from home and digitalise services in general. However, as time went on and the pandemic endured, the need for more permanent solutions arose, and allowed for better planned solutions that included a stronger security focus. ▸

A prediction from last year was that information security would be given a seat at the "adults table" within companies that didn't already consider information security to be a crucial and worthy part of management, as well as an increase in cyber security insurance. If there has been any situations warranting an increased focus on security and cyber insurance, times of insecurity and disruptions tend to be quite high up there.

As predicted, 2020 was a critical year in the advancement of deepfakes. What started as a trend to edit celebrities into different movies has evolved to a point where it is near impossible for the average person to discern the difference from the original, and requiring technology to detect the forgery. Having seen The Queen of England, Barack Obama, Nancy Pelosi and Donald Trump being meticulously and professionally edited into convincing deepfakes, the tools for disseminating misinformation are at the ready, and will continue to be improved and abused.

Now, let's take out our scrying bowls, animal bones and tarot cards. Dim the lighting, fire up some incense, and dive into the future with me.

Dark web actors for hire

Following the natural development of organised crime and Advanced Persistent Threats (APT), more professionalisation and departmentalisation also leads to a complete black market of shady businesses. Where we have more or less full transparency and legislation to govern the security business, the black hats and criminals run rampant on the dark web. Having their own protected and unsupervised (by the good guys) playing field is neither unique nor new. Pirates had their own ports in the Caribbean, and just like then, needs for specialised services rose.

The dark web is the pirate havens of today, where services and access to forbidden fruits are traded like commodities. To maximise their revenue, adversaries will look into opportunities where they can make a profit from other adversaries wanting to exploit a shared or common target. Selling off access or data they no longer need could provide additional business for adversaries, and increase the complexity of a breach.

The premise itself is not new. We have observed cases where APTs, after completing their own objectives on a compromised target, have sold their access and foothold to the highest bidder. Likewise there are other observed cases where after a compromise, APTs did not complete any objectives themselves, but immediately transferred their access. One speculation is the latter implies a business-oriented focus, where the value of their compromised victim is worth more for sale than the chance of the ATP attempting to complete their own objectives and potentially be discovered.

This could also be reversed, where criminal groups specialise in breaching a target, and hand over access to whomever ordered the attacks. There are groups willing to execute their services for a fee. Having multiple, separate contractors executing a breach could also throw off incident responders who would pick up artifacts and signatures from more than one attacker, and may be more effective than operating under a false flag.

APTs offering "breach-as-a-service" adds a new dimension to an already complex landscape for defenders. This growing trend

will enable less sophisticated threat actors to compromise their targets – for a price – and also allow APTs to profit from effectively any organisation they compromise, even after they have completed their own objectives. What is one person's trash is another's treasure.

Increase in attacks against home office devices

One of the more prominent changes in daily work life has come as a direct consequence of the pandemic. We are now confined to our homes, apartments and houses.

The home office has come to stay, and the fear of dropping productivity as a result of people working from home has been put to shame. The collective workforce has proved that home office is a viable option, and many companies are indicating they plan for more leniency for employees to work from home post-pandemic.

This does however mean that you can't rely on physical boundaries for securing your network anymore. The days where you could have your employees sit at their desk and connect their computers to the internal network, and then treat everything on the outside as hostile and everything on the inside as benign are long gone. If the move to the cloud was the death sentence to the old perimeter security model, the move to home office was the executioner.

The rapid shift to remote workers introduced no shortage of challenges for security teams, including a rapid change in user behaviour and habits, and technically solving the unplanned, massive increase in remote users in an incredibly short period of time. For those employees not already equipped to work from home, some companies chose to issue laptops and computers, some chose to let employees take their office computers home, and some adopted a bring-your-own-device (BYOD) model.

As anyone who is working from home office can likely attest to, the boundary between work life and home life starts to become skewed and blurred. Whether or not it is permitted by policy, users will often use their work devices for personal or otherwise non-work related tasks. A reasonable expectation is that this activity will increase as the home-work boundary becomes more blurred.

Security awareness amongst employees is challenging at the best of times, and simply impossible when devices that make up the home office ecosystem are being used by family members. Home routers, IoT devices, and other family members' devices themselves have expanded the available attack surface, and adversaries certainly know this as well. As a consequence of home office being normalised and expected, we predict a rise in attacks utilising home equipment as a vector to gain access to networks and resources.

Misery from misconfigurations

The pandemic ushered in an era of digitalisation, digitisation and digital transformation. Status meetings are being held online, customer interaction via online platforms and general collaboration is expected to be performed in the cloud.

Digitalisation helps modernise services, improving the quality, availability, and ultimately simplifying (or at least intending to) the lives of those consuming the services. There's an inverse effect to this simplification though, which often creates added complexity for those responsible for delivering the services. It is only natural that this added complexity will lead to mistakes and oversights – like misconfigurations.

Misconfigurations are one of the most exploited vectors in the wild, and do not require a significant amount of skill to exploit. Misconfigurations can also be difficult for IT and security teams to discover. While vulnerabilities are a fault or flaw in a piece of software itself, and in the majority of cases can be identified by the version of a software that is running, misconfigurations differ. They are a discrepancy between the intended use (whether defined by an organisation or the vendor) and the actual implementation. So depending on their intentions and policy, the same configuration at one company may be a misconfiguration at another.

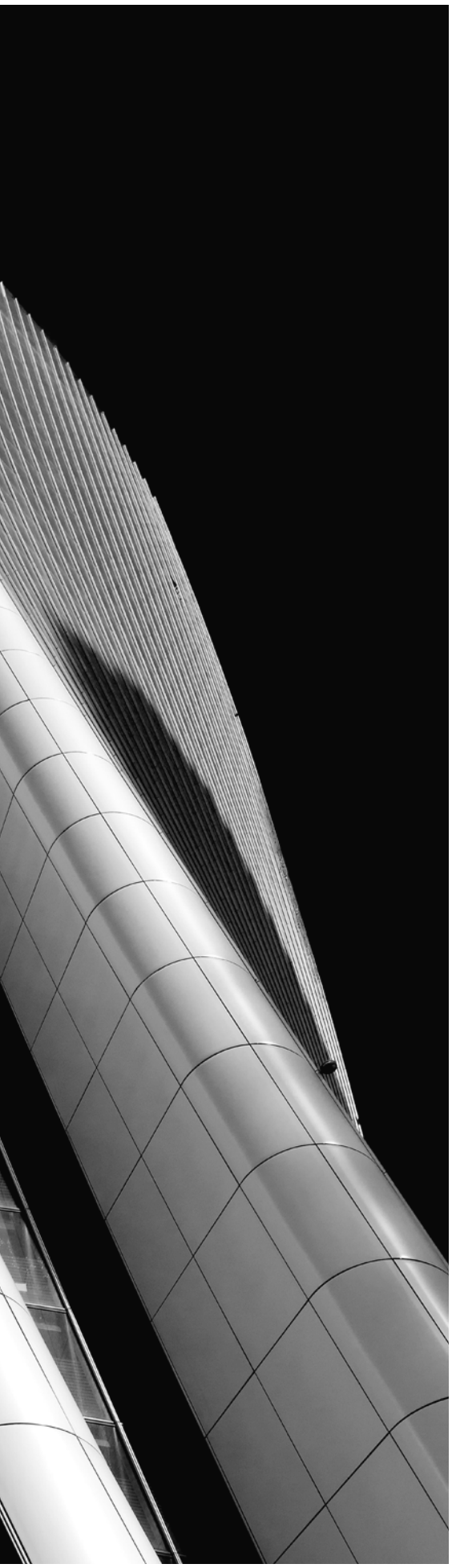
In a basic example, last November, we saw defence ministers in the European Union having their top-secret web meetings gate-crashed by Dutch journalists, indicating that something wasn't configured as intended. Had this misconfiguration been discovered and exploited by someone with more sinister intentions, the outcome may have turned from public embarrassment to an international political crisis.

As the world around us becomes more digitised and integrated, the complexity to manage this ecosystem will continue to grow, and so too will the number and severity of breaches due to misconfigurations.

Multi-channel ransomware goes primetime

Ransomware is just something we have learned to live with.
For adversaries, it's understandable why they continue to focus time and resources on ransomware – it's low effort, high reward, and it works. But that doesn't mean cyber criminals are standing still, and neither are defenders. New detection technology, better backup routines, improved access control, established incident response plans and general user awareness are all contributing to attackers needing to adapt and evolve not only their technology, but their tactics and techniques as well.

For instance after infection, and even before, cyber criminals are directly emailing victims to extort them, and in some cases even calling them. Depending on the corporation or ▸



individual being targeted, there is an array of unpleasant methods and threats to be used to add seriousness to the extortion. Could information about business critical and/or confidential information be distributed? Personal information about specific employees? Information on customers? Threats to shut down critical services? Utilising several channels to follow up the attack has evolved ransomware from being a passive distribution of destructive malware with the hopes of victims paying a ransom, to a targeted, multi-channel attack with a human element on both sides.

A few years ago, the extortion in ransomware was primarily targeted towards the destruction of data – something that was often thwarted by simply restoring the data from a backup (if the backup existed, of course, but that's another topic). Unfortunately these days even the best backup routines will only mitigate one of several potentially devastating consequences created by multi-channel attacks, and this continues to make ransomware a serious threat for every company.

The big picture

Despite 2020 being a year full of surprises and changes, the cyber security scene remained relatively close to expectations. While the who, what and when of cyberattacks are not easily predicted, the fact that they will happen is. There is a continuous cat-and-mouse game between the good guys and the bad guys, and as Baz Luhrmann said in his legendary song "Everybody's Free (To Wear Sunscreen)" back in 1999 (somewhat out of context, but still fitting); "sometimes you're ahead, sometimes you're behind. The race is long". Cybersecurity is an infinite game that has no winners or losers, no beginning or end, and no rules that define the game – only players that enter, change, adapt, and ensure the game is continued to be played.

ARTICLE

Enterprise Security Architecture

Optimise your security investments

André Holvik
Product Manager

Kåre Magne Almåsbakk
Business Development & Contracts

Angel Alonso
Team Leader, GRC

Mark Totton
Principal Management Consultant, GRC

Joakim Kunst
Senior Security Analyst

Per Morten Sandstad
Cyber Threat Intelligence Consultant

Kim Trønnes
Penetration Tester

Tormod Emsell Larsen
Team Leader, Product Architecture

AFTER READING THIS ARTICLE, YOU WILL:

- Better understand how to future-proof your security investments by designing a holistic security architecture drawing on several parts of your organisation
- Learn how to take more informed security investment decisions based on risk, and communicate this in business language
- Get a sneak peek into how we at mnemonic prepare for changes in the threat landscape through our own Enterprise Security Architecture framework

Justifying investments into new security technologies is a challenge for most organisations. Most would want to spend money on relevant risk mitigation initiatives, but at the same time they also need to see a return on their investments. With an increasing demand for larger budgets, management and the overall IT organisation must establish a common ground and a common language to prepare for the future together, in order to find the balance that is just right for your organisation.

To reach this common ground, discussions will often revolve around questions like:

- How much should we invest in cybersecurity solutions?
- How will this new security service help us reduce our overall risk exposure?
- What specific business requirement does it support?
- Are there other smarter and more cost-efficient ways we can spend the money?
- How can we document the need for adding a new technology into our architecture?

Answering these questions is the aim of the concept that is called Enterprise Security Architecture, and this is what we will discuss throughout this article.

Enterprise Security Architecture

Enterprise Security Architecture can be defined as a framework that "describes a structured inter-relationship between the technical and procedural solutions to support the long-term needs of the business" [1].

It is used to prepare organisations for future threats and risks that currently are not on their radar, and ensures investments are fully qualified and agreed upon throughout the organisation. A goal when working on Enterprise Security Architecture is to identify where your organisation has security capability gaps and ensure money is spent where it minimises risk.

At mnemonic, we've worked with Enterprise Security Architecture both for our own use, and through helping our customers ask the right questions and prepare business cases for their security investments. When attempting to solve these challenges for ourselves, we have seen an emerging need for involving all units within mnemonic, be it those working ▸

[1] *See Reference List at the end of the report*

with Governance, Risk and Compliance, System Integration, Technical Risk Services and so on. That is also why you might have noticed that this article has such a wide range of contributors.

Most Enterprise Architecture tools on the market emphasise collaboration and involvement from the overall organisation to meet each unit's need for reporting, metrics, documentation, etc. This includes both collaboration within the team responsible for security investments, as well as between departments to ensure the ideas, designs and practices are adopted across the organisation, also its board.

While there are many mature Enterprise Architecture solutions available, only a few offer extensive feature coverage on the security element within Enterprise Security Architecture.

An Enterprise Security Architecture framework explained

All of the article contributors are part of what we internally call The Enterprise Security Architecture Group. This cross-departmental group was created with a goal of building a framework that enables mnemonic to make future-proof and sound security investments that take the whole organisation into account.

Throughout the process, we've recognised that the lessons learned from what was initially an internal project may benefit other organisations as well. The developed approach seems flexible enough to meet every organisation's individual needs when it comes to building their Enterprise Security Architecture.

The figure below details the core concepts we are basing the mnemonic Enterprise Security Architecture (mESA) framework on.

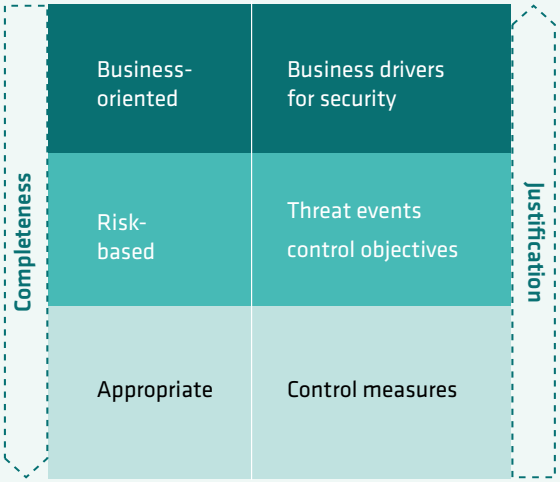


Figure 1: mESA core concepts

mESA incorporates and is inspired by a range of frameworks, security standards, guidelines, best practice and two decades worth of information security experience. Figure 1 illustrates a fundamental principle adopted from SABSA – two-way traceability.

Traceability for completeness: The top-down traceability allows every business requirement to be traced down to the technical controls, and ensures completeness in the Enterprise Security Architecture.

Traceability for justification: The bottom-up traceability, on the other hand, allows every single technical control to be traced back to the business requirements it supports, and ensures business justification for each technical control the organisation invests in.

This traceability makes it possible to identify gaps and risks in the Enterprise Security Architecture, and it provides a way of identifying elements that are not supporting the business requirements and therefor might be unnecessary.

This way, the framework connects the business aspects and the technical elements, and helps us to optimise security investments.

Building blocks

When creating our framework, we found these frameworks, guidelines and best practices to be especially useful:

- *MITRE ATT&CK (Adversarial Tactics, Techniques & Common Knowledge)*[2]

Provides the ability to describe and categorise adversary tactics and techniques based on real-world observations. Emerging as a common language that simplifies classification across the industry.

- *SABSA (Sherwood Applied Business Security Architecture)*[3]

A vendor-neutral community-oriented methodology that adds structure to security architecture initiatives.

- *NIST Cybersecurity Framework*[4]

Well-established framework that defines five primary pillars for a successful and holistic cybersecurity program: Identify, Protect, Detect, Respond and Recover.

- *mnemonic's Security Strategy methodology*[5]

A continuous process that aims to describe the major security concerns an organisation faces and a roadmap on how to minimise the related risks.

Security standards

In order to make the framework flexible and extensible to support many different requirements, we found ▸

[2,3,4,5] *See Reference List at the end of the report*

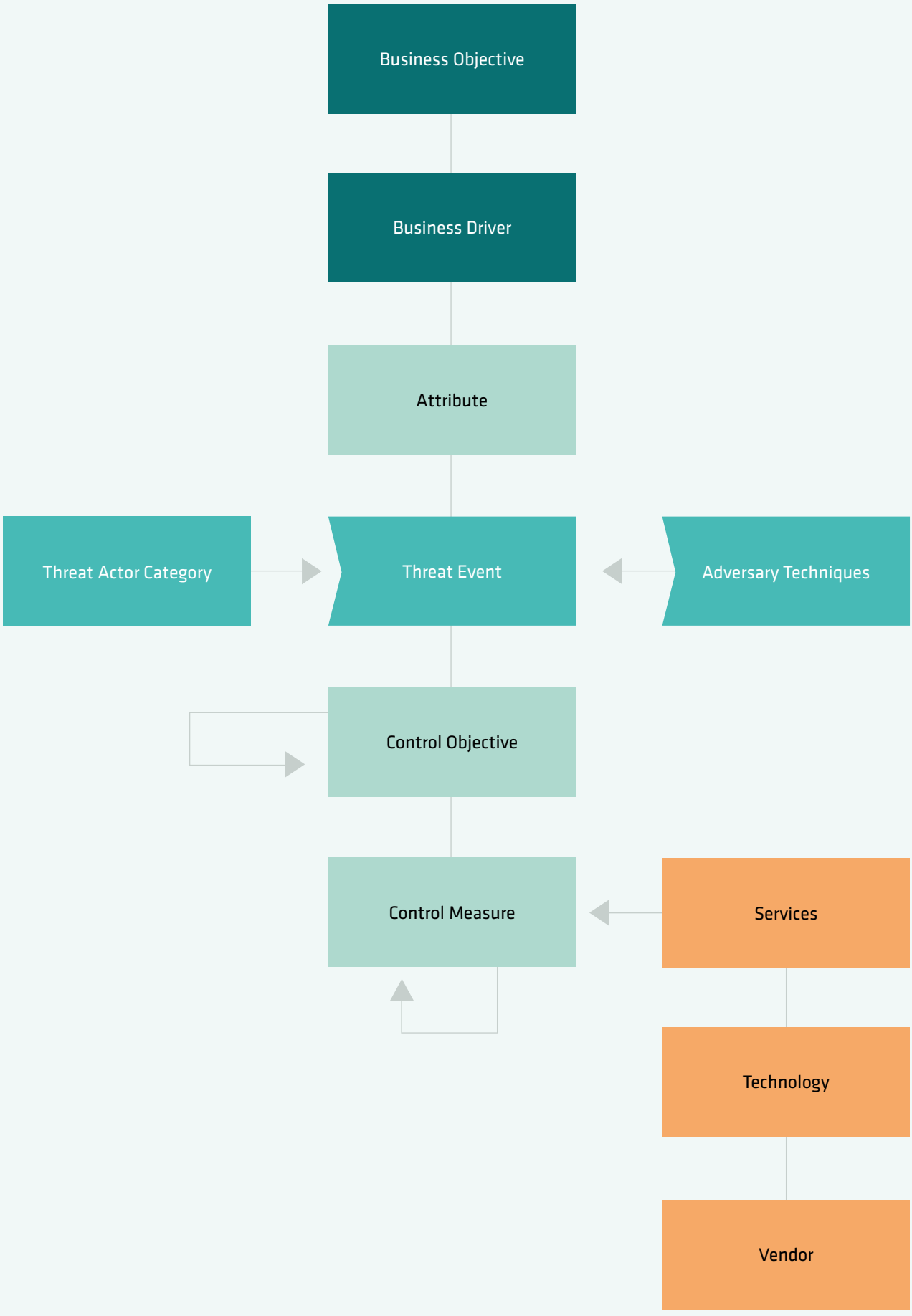


Figure 2: mESA metamodel

it useful to map together several controls from different security standards. At the moment of writing we have incorporated:

- *ISO/IEC 27001*[6]

Provides normative requirements for the development and operation of an ISMS. This includes a set of controls for the control and mitigation of the risks associated with the information assets the organisation seeks to protect by operating its ISMS.

- *CIS Controls*[7]

A set of 20 prioritised key actions that organisations can implement to mitigate known cyber-attacks. They are designed to be implemented, enforced and measured with primarily automatic means. The controls are also known as CIS CSC, CIS 20 and SANS Top 20.

- *CSA Cloud Control Matrix*[8]

The CSA Cloud Controls Matrix (CCM) is a cybersecurity control framework for cloud computing, composed of 133 control objectives that are structured in 16 domains covering all key aspects of the cloud technology.

- *NIST SP 800-53*[9]

Catalogue of security and privacy controls for federal information systems and organisations, and a process for selecting controls to protect organisational operations.

- *The Norwegian National Security Authority's (NSM) basic principles for ICT security (NSMs grunnprinsipper for IKT-sikkerhet)*[10]

Defines a set of principles and underlying measures to protect information systems (hardware, software and associated infrastructure), data and the services they provide against unauthorised access, damage or misuse.

An Enterprise Security Architecture framework in practice

In practice, there are many ways to build an Enterprise Security Architecture. Here we'd like to share how we ended up building our framework:

Business Objectives

The framework begins with business objectives. We need to understand the organisation's business objectives in order to be able to support them. What is our mission? What are our strategic, tactical and operational business objectives?

Some examples include: Be the leading brand, deliver shareholder value, deliver innovative services.

Business Drivers

We collect these business objectives and abstract them into meaningful business drivers for security in business language. We can see these as the requirements we need to support and enable business, as well as bringing demonstrable value to our security initiatives. Their main objectives are to break the language barrier between business and security.

Some examples include: Build trust in our business, build organisational resilience, exploit opportunities in new technologies.

Attributes

We achieve the two-way traceability introduced earlier by enabling abstraction from business context to a business security context. To populate the missing link between business and security we model these business drivers for security into a set of normalised, reusable attributes understandable for stakeholders at all levels.

Some examples include: Available, Confidential, Integrity assured, Private, Recoverable.

Threat Events

Several threat events can affect our attributes in a way that damage our capability to support our business requirement and therefore we could fail to enable our business objectives. An example include: Ransomware. This is one of the most common threat events that organisations face today. Or as it is defined in our framework: "Financial gain by interactive deployment of targeted ransomware".

Threat Actor Categories and Adversary Techniques

We use our own taxonomy to define threat events based on our own Threat Intelligence:

- Each event is related to a set of threat actor categories.

Some examples include: Crime-syndicates, Nation states, Insiders, Sensationalists.

- Each event is associated with several adversary techniques, as defined by MITRE ATT&CK

Example: T1190 - Exploit Public-Facing Application.

Control Objectives

As the next traceability level in the model, threat events and adversary techniques are associated with control objectives to mitigate their damaging effect towards the attributes. We defined control objectives according to NIST Cyber Security Framework subcategories.

Example: DE.CM-4 – Malicious code is detected.

Instead of a pure compliance exercise, we take a more risk-based approach by connecting threats and control objectives. This way, we can focus our efforts on the control objectives that specifically mitigate the given threat, reducing its probability and therefore reducing the overall risk exposure. ▸

[6,7,8,9,10] *See Reference List at the end of the report*

Control Measures

Control objectives are further detailed as specific control measures. We have built mappings to the most popularly adopted security standards in the industry (see security standards section above for a list). This gives us the opportunity to adapt the framework to the preferred implementation standard in each case and to reuse controls across standards. Control measures could be further broken down to physical and component architectural artifacts in a specific organisational context. These can be seen in Figure 3 found on the next page.

Services

We have connected the control measure to all the services that we provide. This allows us to trace how we can best and more efficiently help reduce risks related to the threat events we and our customers are most worried about and that can cause the most damage to our/their business objectives.

Technologies

To optimise our service portfolio across departments we have also connected our services to the different technologies that enable them, creating a consistent technology portfolio. In this way, we can keep our technology portfolio up to date to protect against new threats, and identify additional solutions to meet changing customer demands.

Vendors

Finally, we have mapped these technologies to the different vendors that can provide them, allowing us to streamline and identify new opportunities for partnerships.

Final remarks

Applying detection, control and countermeasures to protect against the ever-changing threat landscape can be challenging. With many frameworks and standards, organisations might end up with too many technical controls and no clear way of prioritising them. You then face the daunting challenge of finding tools and services to address these controls. Ideally, you are now looking for something that can manage multiple controls, but even then you might find that you need more than a hundred tools and services to cover all of the controls.

Few companies have the human resources and budget to handle such a technology stack and would have to prioritise what controls to implement. Our Enterprise Security Architecture framework aims to address this problem by incorporating business objectives and mapping them to the threat that is most relevant to the business. This will provide guidance on what security controls are most relevant for an organisation, allowing it to focus its talent and spending on the challenges that are most important.

We have learned a lot while implementing this framework internally, and we believe the approach can assist others in building their own robust and flexible platforms to help them better prepare for the future. ●

Feel free to visit **https://www.mnemonic.no/mESA** to find additional resources on Enterprise Security Architecture.

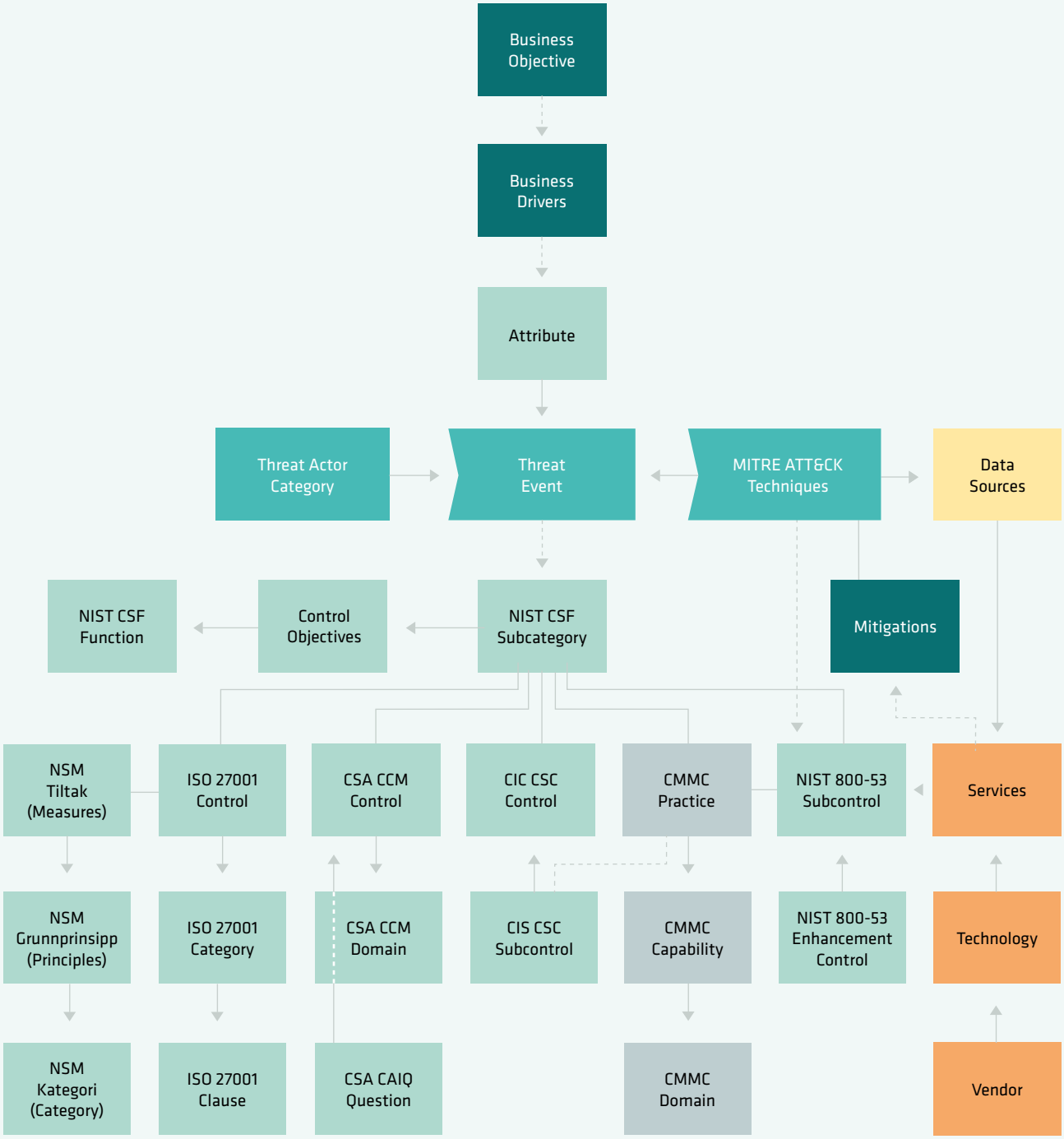


Figure 3: mESA current model with relevant frameworks listed

NATIONAL CYBERCRIME CENTRE (NC3)

Norway

Øystein Andreassen
Police Superintendent

The NC3 is a national centre aiming to prevent, detect and combat threats and crime in cyberspace. The centre provides assistance to police districts and conducts its own cybercrime investigations, in addition to developing the police's cybercrime expertise and methods.

What is your biggest security concern?

Our biggest security concern lately is threat groups utilising ransomware to such an extent that it's leaving large companies and organisations paralysed. The development is worrying, both when you look at the number of cases as well as the consequences this has on society nationally and internationally.

To be clear, ransomware is a serious threat for every company, and can be devastating for both small and large organisations.

In what areas of security do you think we're falling behind?

I'm not sure we would categorise it as falling behind, but a real challenge we and organisations like us are experiencing is that the gap is too large between the threats authorities are warning companies about, and the threat picture perceived by the companies themselves. Hence, many aren't prepared enough for what's coming their way.

This is however, to be expected. Cyber threats aren't as top of mind for those working in a completely difference area, only utilising IT as a tool to meet a very different set of goals. However, it poses a real threat should something go wrong.

What gives you hope for the future of security?

We find hope in the fact that we've been able to develop effective cooperation internationally, as well as across mandates and agencies in Norway. We're also getting further in developing good partnerships and finding new ways of working in public-private initiatives.

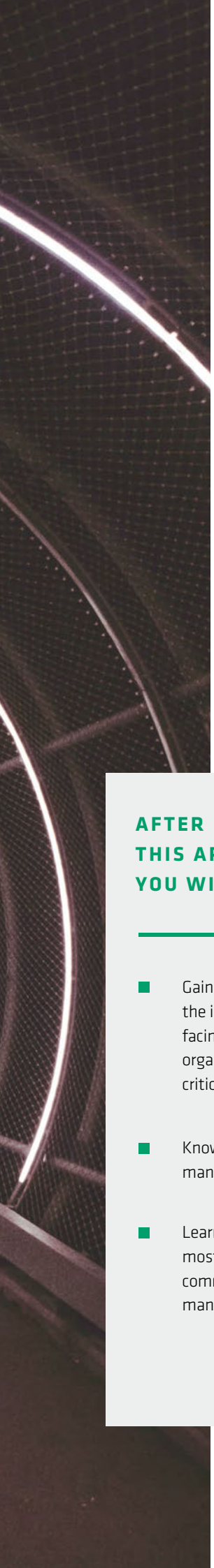
There are many strong forces working together towards the same goal. ●

ARTICLE

We need to talk about insider threats

Anne Aune
Senior GRC Analyst

Kristian Haga
GRC Analyst



AFTER READING THIS ARTICLE, YOU WILL:

- Gain insight into the insider threats facing companies and organisations managing critical assets
- Know how they actually manage these threats
- Learn what they find most difficult when communicating and managing insider threats

A *true agent has instinct, and that cannot be taught. He either has it or he doesn't.* This is a quote from the Netflix series The Spy. The series is based on the life of the Israeli spy Eli Cohen, portrayed by Sasha Baron Cohen. The quote is an accurate reflection of how we typically picture insiders; intelligent moles or spies working undercover. However, real insiders might look quite different from how we're used to seeing them depicted, and as this article will explore, this Hollywood stereotype might actually get in the way of companies and organisations managing real insider threats.

What is an insider?

Insiders can be tricky to define. Not only do they have a range of different methods, motives and backgrounds; the scope and seriousness of their actions differ just as much.

Traditionally, intelligence services have invested a lot of resources recruiting both native and foreign citizens to act as their agents. The goal of these agents is to gain access to valuable information, for instance through extracting information covertly from inside high value targets, like companies and organisations managing critical assets.

However, all insiders are not state sponsored agents. The category also includes people, on their own initiative, stealing assets for personal gains; to benefit financially, seek revenge, or support a personal ideology or political opinion.

Another important aspect when trying to understand what a real insider is, is the fact that insiders don't always knowingly compromise assets. They might not have a clue that they poses an insider threat. Maybe they share too much information in presentations, in dialogue with external organisations they cooperate with or in a social setting. Therefore, research tends to separate between intentional and unintentional insiders:

- An *unintentional* insider is a current or former employee, contractor, or business partner who has or has had authorised access to an organisation's network, system, or data and who, through action or inaction without malicious intent, unwittingly causes harm to the organisation's assets.

- An *intentional* insider is a current or former employee, contractor, or business partner, or simply just someone that an internal employee trusts based on having shared interests, who wittingly causes harm to an organisation's assets.

Reducing insider risks

Insider risks can be reduced by systematically implementing organisational, personnel and technical security controls. This means that to protect against insiders, organisations must coordinate between different departments at different levels.

To use risk management to decide on the level of adequate measures, one needs to define relevant risk scenarios. However, this is not always an easy exercise since risk scenarios concerning insider threats tend to be seen as farfetched, as most companies and organisations consider it unlikely to have an Edward Snowden is on their payroll. ▸

An important aspect of insider threats, especially the state sponsored type, is that it deviates from other types of security risks when it comes to timeframe. While many other security risks result in quick return on investment for the adversary, we have seen that insiders supported by intelligence services often take much longer before a return on investment is materialised. This means that when analysing the attack pattern, one could expect that the earliest phases, such as reconnaissance and the initial intrusion, may take years.

Such complex attack patterns where malicious activity is masked as normal behaviour over time pose a challenge for the security organisation when deciding on relevant risk scenarios.

For this article, we interviewed several C-level executives in three companies that manage assets that the Norwegian society depends heavily on. All of them were chosen because of the high-value assets they manage, and hence their risk exposure to insider threats.

One of them is subject to the Norwegian Security Act and process classified information, the two other companies deliver services to companies subject to the Norwegian Security Act and also manage services that are critical to society.

Interview findings: three companies managing critical assets

As security consultants, we rarely see insider threats defined and prioritised in companies and organisations' risk analyses, despite most authorities highlighting it as a highly relevant risk both for private and public organisations. We also know that the *real* number of insider cases is much higher than the number reaching the public through the media.

For several years, European law enforcement, agencies and alliances such as ENISA and NATO have reported that the risk of insider threats is particularly high for companies and organisations with critical assets. Some of these are subject to strict national regulations with intrusive mandatory measures, such as security clearance and authorisation regimes. However, many companies and organsiations managing high value assets, such as private companies and research institutions, are not subject to the same regulations.

Our experience is that companies and organisations are struggling to prevent and mitigate insider threats. Therefore, we decided to interview three Norwegian companies that manage critical assets to find out more about the challenges they face.

Do you communicate potential insider threats internally?

Security specialists and authorities tend to recommend insider risks to be reduced by implementing controls in line with a security management system. In any management system, communicating plans and implementations is essential to making the management system operational. One cannot expect an insider risk to be managed and controlled if the reasoning and chosen countermeasures are not communicated internally.

Finding 1: Mitigation of insider incidents is discussed amongst top management. However, the prevention of insider threats is not formally addressed in a broader forum reaching all employees.

All interviewees said that it is challenging to talk about a potential insider in a broader context. Two interviewees report that they have had to discuss insider threats at the top management level as a result of managing actual incidents. They mention the need for an actual case or a proper reason to communicate insider threats in broader terms. This essentially means that insider threats are not proactively communicated, but rather reactive and on an ad-hoc basis if certain situations require it.

We often evaluate the maturity of a security management system from Level 1 to 5.

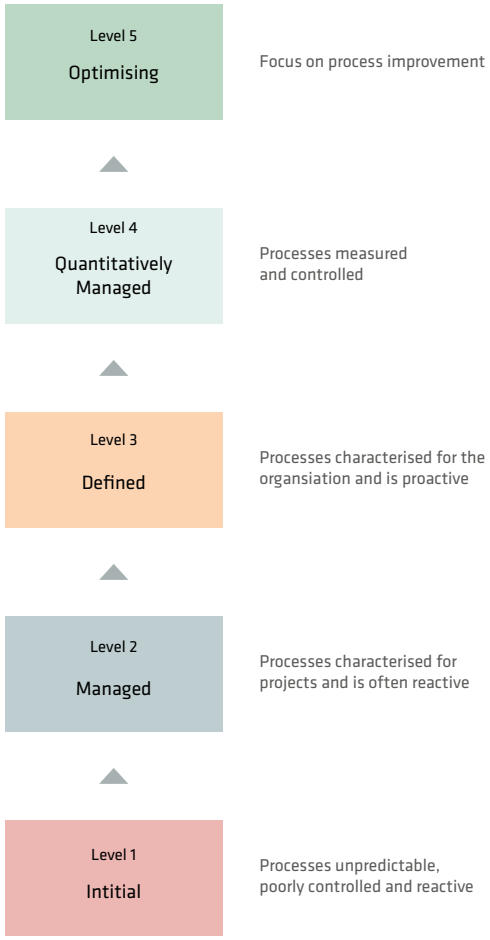


Figure: characteristics of maturity levels inspired by CMMI Institute[1]

[1] *See Reference List at the end of the report*

Our experience shows that many companies and organisations' Information Security Management Systems (ISMS) fall within Levels 2 and 3. However, when it comes to managing insider threats specifically, the maturity is often lower. Regulations like the Norwegian Security Act and GDPR require security controls to be evaluated to ensure they work as intended. Therefore, when assessing companies and organisations managing critical assets, we typically expect them to aim for maturity Level 4 (as a minimum).

If the insider threat is not communicated by top management and not included in risk management and other basic processes in the security management system, the maturity level can at best be expected to be at Level 2.

Finding 2: Insider threats are not formally covered in meetings with employees without management responsibilities

Both unintentional and intentional insider threats are relevant for employees at all levels within an organisation. None of the interviewees said that their company had formally communicated insider threats in a meeting with a larger number of employees. The sole exception to this was one organisation that had included some aspects of insider threats in their e-learning program that was delivered by an external vendor.

The interviewees elaborated on reasons why it is difficult to talk about insider threats with employees. A common theme was that it was difficult to find the right balance between trust and control."We are a company based on trust" and "we must trust our employees" are comments all interviewees mentioned. Together with statements like:
"We cannot risk creating an environment that stigmatises employees or foster a culture of suspicion."

An important takeaway from these companies is that focusing on the assets that need to be protected, made it easier to communicate why they needed to take actions against insider threats. As an example, one interviewee mentioned that employees at a large bank tend to understand the need for securing the bank's assets because they know what value the bank possesses; not because they think their colleagues are thieves. Hence, a first step to be able to communicate insider threats and your chosen countermeasures is to make sure the employees understand the assets they manage, and the value of these assets.

Do you address both intentional and unintentional insider threats?

A lot of research, and not to mention the media, focus on the intentional insider – the spy. But what about the employee discussing internal information with someone at a conference? Or when someone includes detailed information in an external presentation to impress the audience? ▸

Or they become the life of the party by talking about information very few know about? Amongst the interviewed companies, two had mainly considered the intentional, malicious, criminal employee who leaks information to harm the company, in their discussions about insiders, or in their security management system. The third company, who is subject to the Norwegian Security Act, had spent some time on the unintentional insider. They mentioned plans of launching a targeted awareness campaign on the topic of unintentional insider threats towards exposed employees, based on their role or specific department.

Finding 3: When communicating to a broader audience of employees, it is easier to speak about unintentional insiders

All interviewees confirmed that it is much easier to communicate insider threats to a broader audience of employees if focusing on the unintentional insider rather than the intentional one. Presumably this is, at least in part, because the unintentional insider scenarios can be associated with honest mistakes that can happen to anyone. This is contrary to the intentional insider that is making a decision with intent, and thereby brings the ethics, values and trust of employees into question.

How do you communicate an insider incident externally?

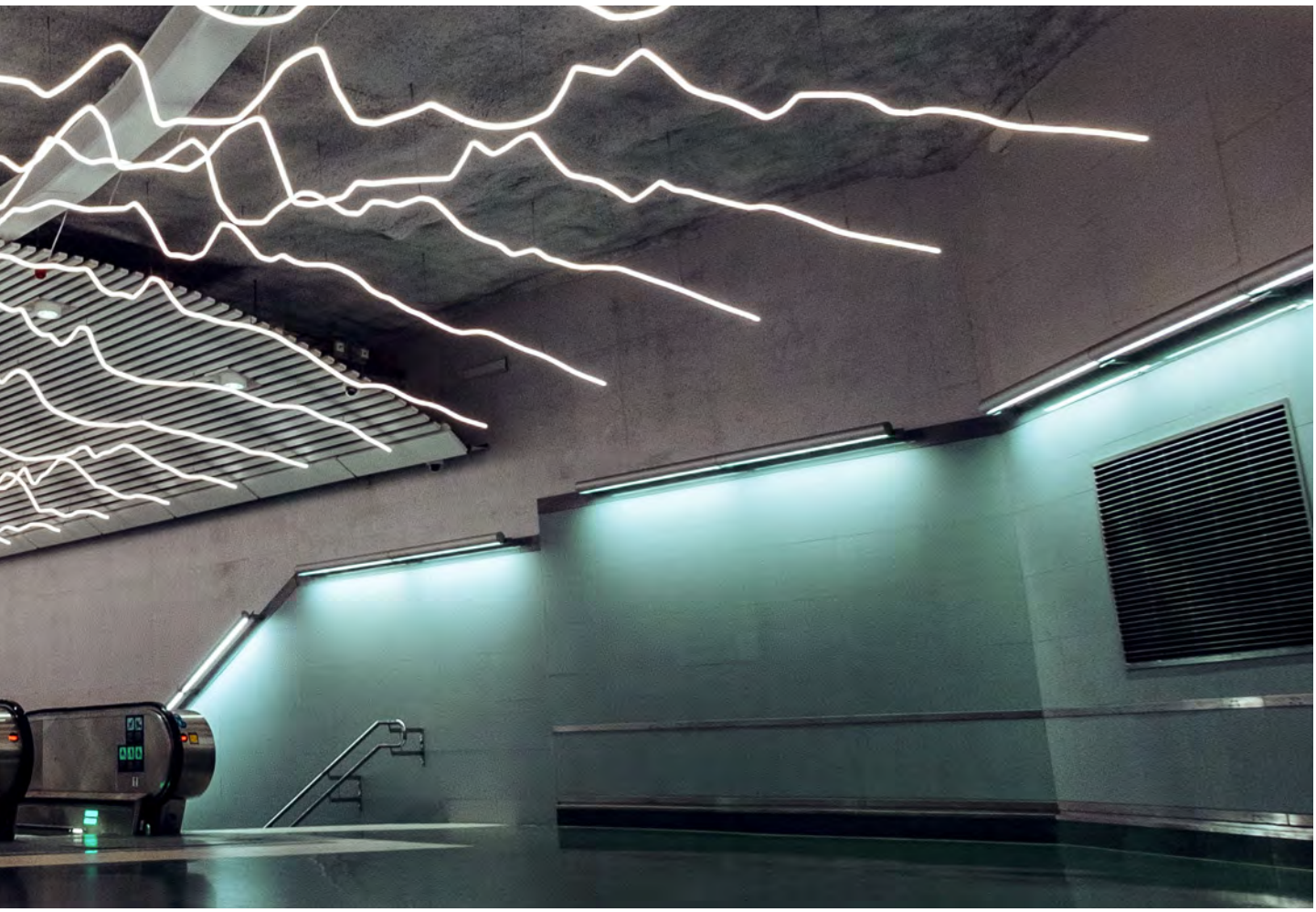
Most insider incidents are never communicated externally. But for some organisations, this type of incident leaves them with no choice. Customers, investors, partners, regulators and other key stakeholders will need to be informed, and by proxy, the media.

The interviewees are concerned that a security incident, and especially one with an insider, would destroy the company's reputation.

As part of our research, we spoke with a media spokesperson in one of the companies that had experience in communicating an insider incident to the media.

The spokesperson's position was that there are two outcomes for a company choosing to externally communicate an insider incident: the way you handle the situation will either increase or decrease the trust people have in your organisation.

Finding 4: Saving the company's reputation is the first priority when considering how an insider incident should be communicated externally



The chosen strategies of the spokesperson to increase the likelihood of gaining trust were; being *active*, *open* and *honest*. The spokesperson remained active by answering media inquiries, and not saying no to interviews, while also focusing on being open and honest by not withholding information. The spokesperson believed that withholding information would result in the company not owning their own story, making journalists and the public speculate on alternative explanations, which in turn could prolong the media focus.

The interviewees mentioned that they would benefit from more experience sharing when it comes to incidents. Hence, if more companies and organisations followed the strategies mentioned by this spokesperson, others could in turn become better prepared, reduce their risks and actually handle insider incidents better. However, it is worth mentioning that some organisations with classified information and assets cannot publicly share their experiences.

Do you know whether your controls are adequate?

All of the companies interviewed have a security management system in place.

All of the companies interviewed have measures implemented in the recruitment process.

All of the companies interviewed have an incident management process and a system to report incidents.

All of the companies interviewed have security training and e-learning for all employees.

Is this enough?

It's a difficult question. Despite having these controls in place, two of the companies reported that they have experienced insider incidents in recent years and stated that in general they feel unsure whether they are doing enough. What else needs to be done in addition to the general security measures that the companies already have in place?

The Norwegian National Security Authority (NSM) released a new guidance on personnel security in August 2020[1].

The Norwegian National Security Authority (NSM) is a cross-sectoral professional and supervisory authority within the protective security services in Norway. Their areas of responsibility include cyber, personnel and physical security.

[2] *See Reference List at the end of the report*

This guide targets all companies and organisations in Norway, not solely those subject to the Norwegian Security Act. Often such guidelines are more general regarding what measures they propose. NSM's guide however, suggests targeted measures for reducing the risk of insiders.

Some of the measures suggested in this guideline can be considered effective at reducing insider threats, but at the same time intrusive to privacy. Systematically gathering sensitive information in meetings with employees and identifying changes in personnel behaviour are examples of such measures. Both of these measures are commonly implemented in companies and organisations with classified information (hence subject to the Norwegian Security Act), for instance in the defence industry or certain government agencies. However, these measures are not commonly found amongst those that are not subject to such regulations. Can companies and organisations managing critical assets, but without being subject to the Norwegian Security Act, actually implement such measures?

Finding 5: There is uncertainty around whether intrusive measures are legal

The company we interviewed that is subject to the Norwegian Security Act already had procedures in place to perform vulnerability assessments of employees. This is mainly conducted through the authorisation regime imposed by the Norwegian Security Act. However, the same company mentioned that the majority of their employees are not security cleared and it is difficult by law to ask employees about potential issues regarding their background, financial status, addictions, close relatives, and other questions that are covered through security clearance and authorisation. This concern was echoed in the other interviews.

None of the companies systematically assessed personnel behaviour, and all three mentioned that they did not understand how to conduct such an assessment.

GDPR mandates a lawful basis for intrusive privacy measures. When considering a vulnerability assessment in the form of a meeting between the employee and the manager, one should evaluate the lawful basis. *Consent* cannot be used as a lawful basis for processing because of the power imbalance between the employer and employee. Hence consent cannot be "freely given" in this setting. *Legal obligation* can be used by some, for example those subject to the Norwegian Security Act. However, for other companies and organisations, measures are left to be justified by *legitimate interest*. This means that the employer needs to demonstrate that its own interests clearly overrides the interests or fundamental rights of the employee.

Another justification could be that intrusive privacy measures are necessary when companies and organisations preform a task in the public interest. However, this is not something they should decide for themselves. The mandate to decide on using this as a lawful basis is likely to be delegated to an official authority.

Understandably, this leaves many uncertain. If they manage critical assets on behalf of society, do they have a legitimate interest to implement intrusive measures to reduce insider risks? The guideline by NSM makes a reservation that companies and organisations themselves must assess the legal basis for implementing the proposed measures.

Final remarks: reducing the uncertainty

To conclude, this assessment shows that companies and organisations are:

- Having reactive efforts, but are trying to be more proactive and are looking for information about how other comparable companies and organisations are managing insider threats.
- Not communicating the risk of insiders to a broader audience internally and are concerned about the effects it can have on company culture.
- Showing a clear interest and willingness to manage insider risks, but are uncertain about what constitutes adequate and legal measures.

There is a need to improve the maturity of how insider threats are managed, by taking the step from being reactive to proactive. An essential part of being proactive is to project the risk, which is difficult and probably not possible if management don't talk about the risks associated with insiders.

When it comes to strategies for communicating insider threats, it's beneficial to include unintentional insiders in internal communication, as well as focusing on the values of the assets managed by the organisation. This will make it easier to effectively communicate insider threats, and in turn manage insider risks.

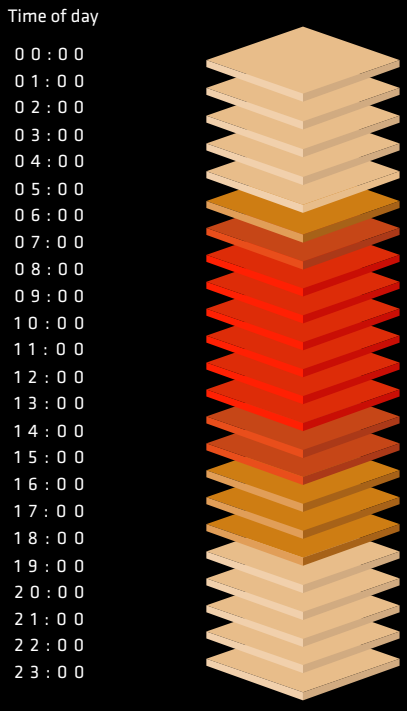
In order to effectively enable companies and organisations to manage their insider risks, the uncertainty must be reduced. As a potential starting point, it would be beneficial if authorities and relevant agencies specified the types of organisations and companies that should evaluate whether they have legitimate interest for implementing more intrusive measures.

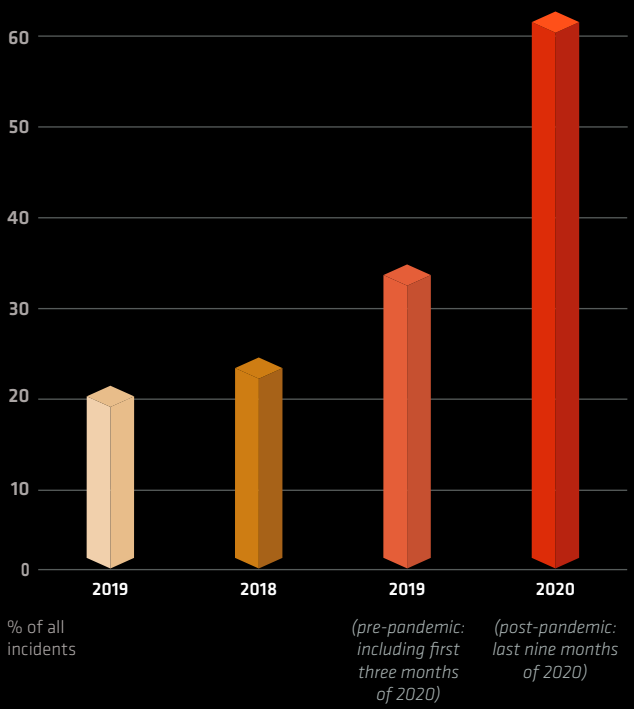
2020 - A VIEW FROM MNEMONIC'S SECURITY OPERATIONS CENTRE

All statistics are collected from the analysis of nearly 7 trillion security events and over 38 000 real customer cases detected in our Security Operations Centre

WHEN ARE SECURITY INCIDENTS HAPPENING?

Security incidents continue to happen at all hours of the day. 71% of security incidents are concentrated between the peak working hours of 07 – 15. This supports the generally accepted truth within the security community that more user activity tends to lead to more security incidents - or put differently, there's a direct correlation between user activity and security incidents. The volume of incidents consistently peaks in the morning when users are first logging on, and often quickly working through their collection of email from the previous evening. Our observations have repeatedly shown that users are more prone to inadvertently clicking malicious links, opening hostile attachments or visiting suspicious websites when they are tired, hungry, or likely to be paying less attention to individual emails, such as clearing their inbox first thing in the morning.





ABUSING IDENTITIES SKYROCKETS TO THE CLOUDS

There is a rising trend in security incidents where a user or attacker has gained unauthorised access to some resource. Most commonly this is through the abuse of user and administrator credentials. The trend has been growing somewhat steadily from 2017 to 2019, and is connected to the continued adoption of cloud services. As identity becomes the new perimeter, attackers continue to see increasing success in abusing credentials in cloud and remote services. This trend was exemplified towards the end of March 2020 when a large portion of the world's digital workforce entered lockdown, and many were forced to work from home.

There was a surge in cloud and remote work tools being implemented and used, and the attackers took notice. In 2017, 20% of all security incidents involved this type of attacker behaviour, while increasing to 23% in 2018. If we consider 2019 to be our pre-pandemic baseline and include the first three months of 2020 to the figure, the trend rose to 34%. Post-pandemic, we saw this figure skyrocket to an astonishing 61% of all security incidents involving some type of account abuse.

2020 saw an

83%

increase in reconnaissance activity, and has grown nearly 5x since 2018

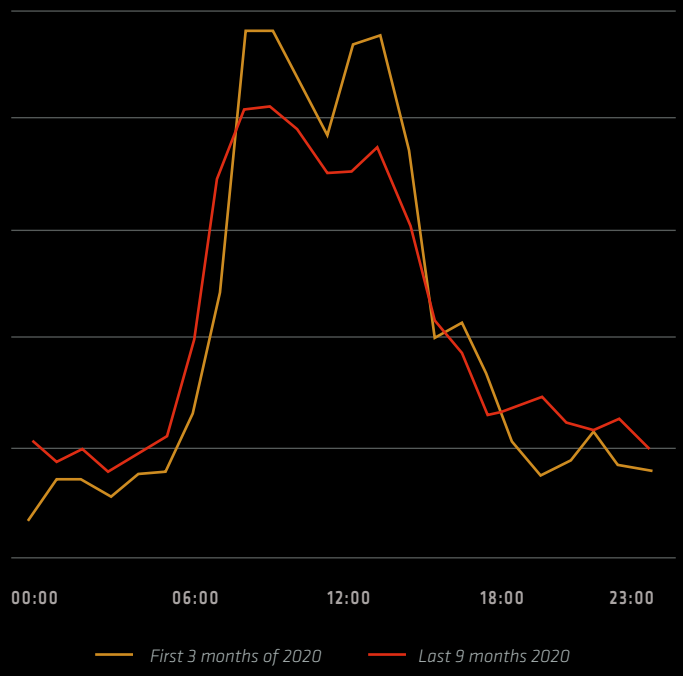
INCREASE IN RECONNAISSANCE

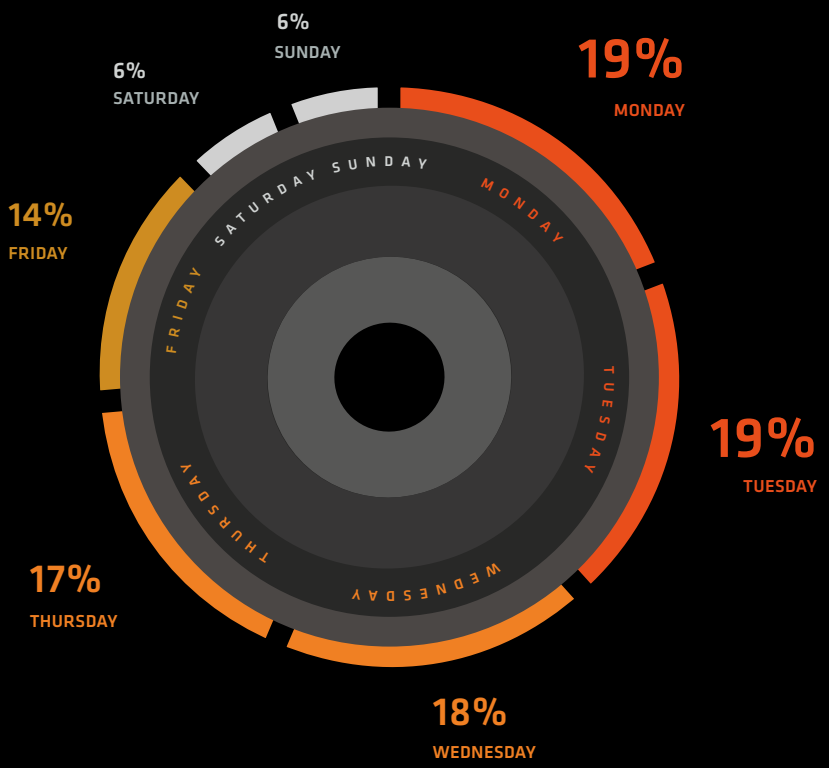
In 2020, we saw NetLogon, SMBGhost, the SolarWinds compromise, the pandemic as a whole, and other vulnerability disclosures that led to noticeable spikes in reconnaissance activity and active exploit attempts against those vulnerabilities (or in the case of the pandemic, social circumstances). These spikes often occur in the first two weeks after an exploit is made public, which reduce over time until they become part of the collective scanning and probing noise that plagues the Internet.

THE IMPACT OF A REMOTE WORKFORCE ON SECURITY INCIDENT TRENDS

The hourly trends from 2019 and 2020 look very similar to each other. However when comparing the first three months of 2020 to the last nine months (pre and post covid), we can see in the latter portion of the year that more incidents are occurring after hours, and the increase and decrease on either end of the work day is more gradual – or if you would, the curve is flattened. Our speculation is that once the majority of the digital workforce began working from home sometime in March, their working hours and habits changed. Between dealing with closed schools, limited daycare hours, lockdowns and other schedule changes, along with an increased ability to work remotely, more people were starting their day earlier, working later and in general changing their habits.

Historically we have observed a noticeable dip in security incidents during the lunch hour, and an increase after lunch as people log back on. This is quite noticeable in the first three months of 2020. However once people started working from home, the lunch dip is far more gradual, and there is no significant peak or increase after lunch. One explanation may be that when having home office, more people are either taking shorter lunch breaks or simply working through their break, thereby reducing the 'logging on' effect.





WHAT DAYS ARE SECURITY INCIDENTS HAPPENING?

Security incidents occur every day of the week, though as expected, there is a significant increase on weekdays. The decrease in incidents occurring on Fridays can most probably be attributed to users working less frequently on these days – namely due to public holidays and users taking long weekends.

ARTICLE

Gatekeeping

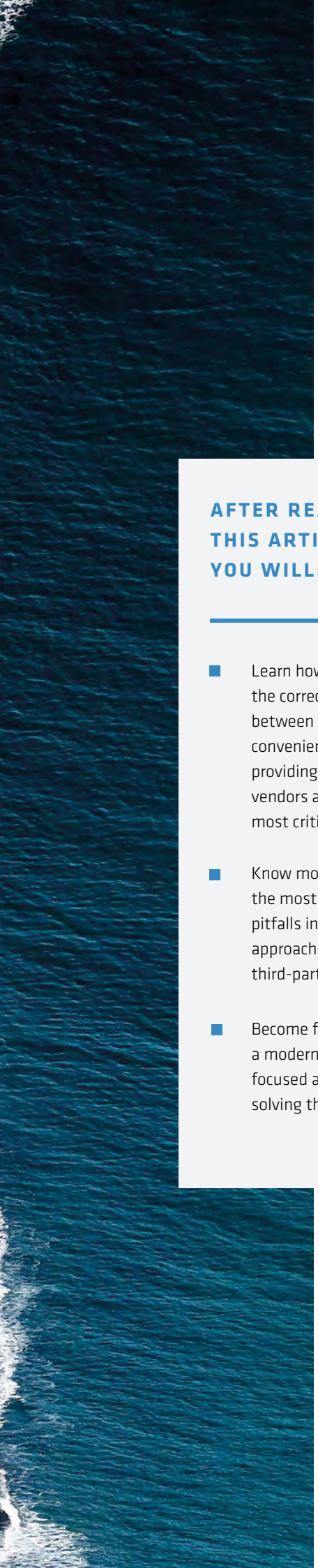
Shining a light on unsanctioned remote and third-party access practices

Adrian Helle

Security Infrastructure Specialist

Konrad Halnum

Security Infrastructure Specialist



AFTER READING THIS ARTICLE, YOU WILL:

- Learn how to find the correct balance between security and convenience when providing third-party vendors access to your most critical systems
- Know more about the most common pitfalls in traditional approaches providing third-party access
- Become familiar with a modern, security-focused approach to solving this challenge

Meet Bill. Bill is responsible for securing the control systems in a large hydroelectric power plant. Bill's company performs work that could be classified as critical infrastructure by the authorities, and the control systems Bill is responsible for manage equipment that is critical to the operation of the company. Bill has a lot on his plate.

Like most control systems, these will require regular maintenance. In order to perform maintenance on proprietary and complex systems, Bill will typically request assistance from external specialists, sub-contractors and the control system vendor themselves. These specialists need to access the systems for tasks like management, maintenance and support.

This raises some important questions for Bill:

- How do I secure the access to these systems, and ensure compliance when my most critical machines and software are opened to external sub-contractors?
- How do I keep track of what the contractors are doing in my systems, intentionally or not? Which commands are they running? Are they leaving any software behind? Are they unknowingly using a compromised client that could create a bridgehead for threat actors into my most critical systems?
- How can I document my routines and provide a complete audit trail to stakeholders, including government entities and regulators?

Does Bill's situation sound familiar? Our experience is that most organisations of a certain size, critical infrastructure or not, will ask themselves these questions when they use third-party vendors. This article will explore this specific use case and suggest a potential pathway for answering Bill's concerns.

Common ways to provide third-party access

Finding a secure way to provide third-party access for organisations with a control system environment is not a new challenge, and there have been many attempts at solving this in the past.

First, let's take a step back and look at the security zones you commonly find in most organisations' networks. These typically include:

- various cloud services, for instance providing video conferencing and Office 365
- internet exposed services and systems in the on-premise datacentre
- internal systems such as a development environment, and data storage on the local plant networks
- the industrial control systems (ICS) located on the process network, our most critical zone ▸

Most organisations would want to maintain a strict access control policy to the secure zone below the process network seen at the bottom of Figure 1, whilst still allowing for necessary maintenance. This often results in strict policies which are difficult to implement and use. Experience shows that when security controls are too intrusive or impact efficiency, those working with Bill's systems will create shortcuts with undocumented openings and ways of storing data in order to "get things done". These shortcuts can be made with the best of intentions, such as efficiency, but they are still outside the ownership and control of Bill's department, and hence pose a severe security risk. These shortcuts are often called *shadow IT.*

We've seen a wide array of shadow IT examples that make people like Bill want to tear their hair out:

- unsanctioned ADSL lines connected directly to secure zones for remote access, bypassing the security measures in place
- TeamViewer-sessions left open to private PCs
- permanent SSH-tunnels from secure networks into clients on non-secure networks
- temporary storage of data, such as configuration files containing keys and secrets, in unsanctioned cloud solutions like Dropbox or Pastebin

Due to the undocumented nature of these workarounds, if by chance your company network or data is compromised, the nightmare is complete. Without any visibility into solutions deployed by shadow IT, the ability to perform forensics is severely hampered. In other words, if Bill enforces overly strict policies and procedures designed to keep his systems safe, it may in itself pose a risk to his company's operations and security.

Virtual Private Networks

A common way to grant service providers access to ICS controllers and servers is through a static, often unmonitored, VPN connection. This is problematic since the service provider now has 24/7 access to the network through a client controlled by the provider, and creates an opening for supply chain attacks should the provider or client be compromised. It also typically lacks monitoring and auditing trails, meaning the system owner doesn't have much, if any, visibility into what the VPN connection is being used for. As such, it severely limits the overview and control the owner has over their systems.

Remote Desktop Protocols (RDP)

Another common approach to enable remote access is by having a client in the network close to the ICS controllers and servers configured with RDP, VNC, or even tools like TeamViewer. Because the client generally has to run legacy software in order to work properly and communicate with the control systems, the client is often not part of normal patching routines. In ICS environments, operating systems that have long reached their end-of-life such as Windows 98 and NT4 are not uncommon. The problem introduced by this is having an unpatched client exposed on the internet, again giving a threat actor an easy target.

Additionally, it is not uncommon that these login credentials are static, shared amongst employees at the service provider, and rarely changed. The remote service also tends to be "always listening" for inbound connections, allowing service providers' employees to log in at their own discretion.

Balancing the equation

The aforementioned solutions expose organisations to supply chain attacks if the service provider is compromised, and therefore are not recommended as self-standing solutions. So while convenient for the service provider, and perhaps convenient for the operational staff on Bill's team not having to worry about the service provider's ability to access critical systems for maintenance, there is an imbalance of too much risk versus efficiency. Hence, this is often not permitted by many organisations per policy. This strict policy again leads us back to shadow IT.

In the event of a breach, one can argue that the principally less secure practice of establishing VPNs or enabling remote desktop protocols might still be better than risking shadow IT. The solutions above are at least somewhat documented, with, albeit limited, logs and monitoring of what is happening on your systems. Should an incident occur, undocumented access is almost impossible to detect for both internal and external investigators, especially if the parties involved in establishing this access fails to report its existence. In conclusion, a compromise needs to be negotiated to balance security with usability.

Moving out of the shadows

Bill is aware of these traditional ways of solving third-party access, as well as their respective challenges. In order to avoid the most common pitfalls, be it having undocumented and unmonitored workarounds, limited system overview and control, or unpatched clients exposed on the internet, Bill establishes a few requirements and builds an access control system. His system spins up short-lived, actively monitored virtual machines (VM) that only allow the strict communication necessary for the third-party to perform his or her work.

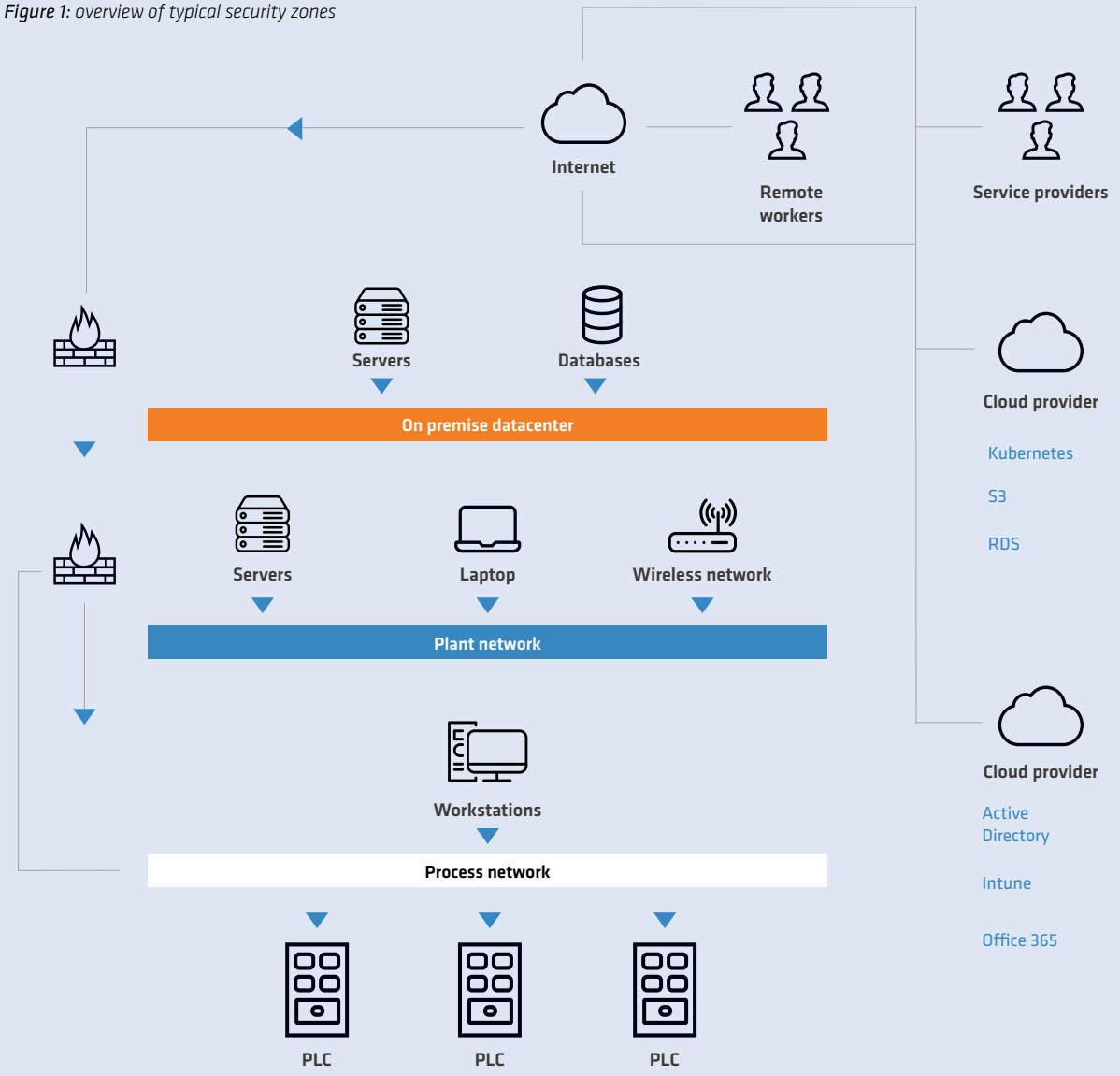
What follows is a discussion and evaluation of how Bill's approach meet his challenges and requirements.

Securing access

Requirement: The system must have a single entry-point which is controlled and owned by the system owner: in this case, Bill.

Sub-contractors that need access to Bill's infrastructure have to submit a request to a ticketing portal. By utilising multi-factor authentication and previously vetted and approved accounts unique for each user, Bill ensures that the user asking for access is authorised.

Figure 1: overview of typical security zones



Upon this verification, a temporary virtual machine is created that the sub-contractor accesses. This VM should only have a limited life span, and takes a fail-close approach that will kill all sessions if any of the following events occur: time runs out, user logs out, or suspicious/malicious activity is detected.

Bill has opted for a solution where all sub-contractors gain access to this VM through a web-GUI. This creates a cross-platform access point and enables Bill to funnel all sub-contractor connectivity through a single point.

By utilising short lived VM's that isolate sessions and removes persistent connections, the sub-contractor or malicious entity cannot re-use the connection. In addition, this approach makes it possible for the connection to always be established with an updated VM through an automatic build pipeline, which is an added convenience and takes patching access clients out of the sub-contractor's hands.

Monitoring

Requirement: Access has to be limited to the task being performed. That means no more continuous, unmonitored access.

By only providing third-parties with a web-portal to access his systems, Bill ensures compatibility and opens for additional opportunities in securing HTTP traffic. The web-portal gives the user either SSH or RDP/VNC access, is easy to use and only requires a web-browser: no configuration of the user's machine needed. This also enables the use of an advanced web application firewall (WAF) /application delivery controller (ADC) to inspect the incoming traffic for malicious use.

An important aspect of Bill's approach is that it also creates an opportunity to perform security monitoring and integrate the monitoring of his industrial environment with his security operations team. Should a malicious actor compromise or gain access through the connection, it can be detected, ▸

and Bill's SOC will be able to assess the situation and terminate any active sessions if necessary. This solves the challenge of unmonitored remote access and Bill gains visibility into what goes on when a sub-contractor accesses internal systems.

User-friendliness

Requirement: User-friendliness has to be a top priority for anyone performing legitimate work in Bill's systems, both service providers and internal employees.

Bill makes user-friendliness a priority so that his users are not tempted to use shadow IT or bypass the controls. His workflow might then look something like this:

- Sub-contractor user creates ticket requesting access to a system for a specific timeframe.
- SOC evaluates request and grants access via web-portal:
  - System owner receives a ticket for auditing purposes.
  - Sub-contractor user receives automatic answer when access is live.
- At the scheduled time, access is granted, and the user can log in. The user is presented with a functional virtual machine that is completely under the control of Bill's organisation.
- From the point of logon until the access is revoked, all activity is monitored by the SOC.
  - If any malicious behaviour is detected, the SOC operator can immediately terminate the session and start incident response activities.
- When the permitted time of access is expired, the user's access is revoked and all sessions are terminated. Revoking of access is logged for auditing purposes.

By providing a user-friendly solution and taking configuration of client and system access out of the sub-contractors' hands, they have fewer reasons to use shadow IT and are encouraged to use the sanctioned access instead.

Documentation

Requirement: Should an incident occur, Bill needs to be able to pinpoint what happened, when and where, and possibly also provide documentation to official government bodies and stakeholders.

Bill knows the importance of documentation and reporting. In his approach, he makes sure:

- all system commands, network activity and authentication data are stored on-premise and made available through a SIEM-like solution
- ready-made reports show every step a sub-contractor has taken during maintenance work
- the solution is system-agnostic, and can be applied in a wide variety of use cases with the same implementation

This is where Bill's ingeniousness really shines through, as these capabilities will provide him with documentation and audit trails.

Final remarks

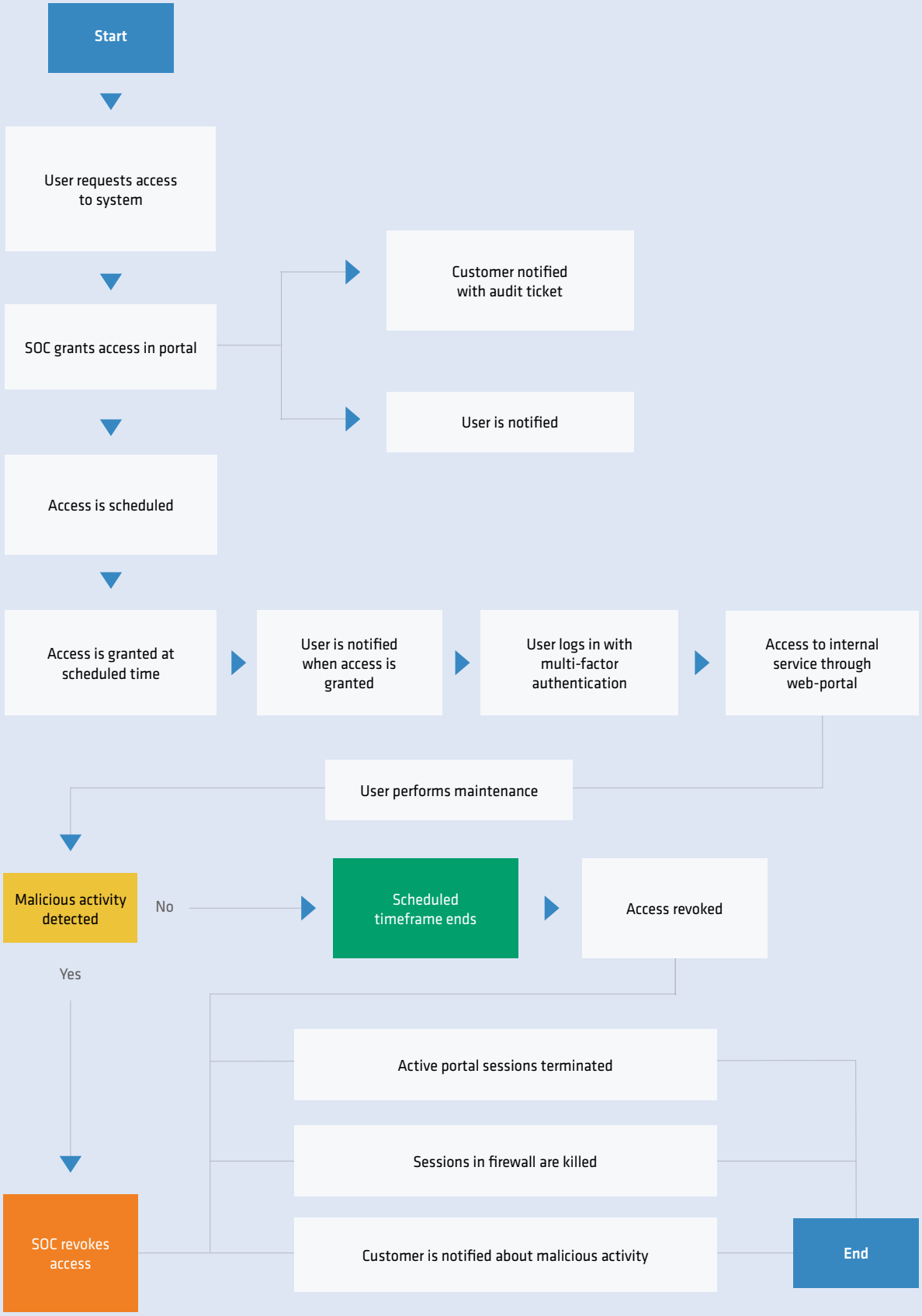
Bill's suggested approach is a modern, security-focused and user centric approach to providing remote access to third-party vendors that may be suitable for many organisations.

If you feel like you can relate to Bill and his challenges, a healthy exercise may be to ask yourself the following:

- Are you confident you have control and visibility into what sub-contractors are doing in your networks and systems?
- Is your process for sub-contractors accessing internal systems to perform maintenance and support user-friendly enough to discourage shadow IT?
- Are you monitoring your sub-contractors for intentional and non-intentional malicious activity?
- Are you keeping records of this activity, for internal and external audits?

If you are not able to confidently answer "yes" to all of these questions, it might be time to re-think the way you are providing access to external parties into your most critical infrastructure.

Figure 2: *: flowchart demonstrating a more secure method for providing remote access*



EYE

The Netherlands

Piet Kerkhofs
CTO

EYE is a Dutch cybersecurity company founded in 2020 with deep roots and extensive experience from the Dutch government's cyber program. EYE provides managed cybersecurity services combined with cyber insurance to European small and medium enterprises (SME), and are passionately involved in making Europe a safer place to do business.

What is your biggest security concern?
Cyber adversaries targeting supply chain companies. We think that SolarWinds is just the tip of the iceberg. This we believe, because supply chain companies are becoming even more of an interesting target for adversaries, in combination with an often underdeveloped security posture amongst these companies.

In what areas of security do you think we're falling behind?
Providing solid, no-nonsense measures for companies that can be considered supply chain targets. These are often SME companies that do not have a 24/7 SOC or an MSSP to protect their infrastructure and data.

We believe that providing this segment with measures like advanced "enterprise" technology like EDR solutions and monitoring during office hours can make a difference. Supply chain companies will then have a more real chance to identify threats before it's too late.

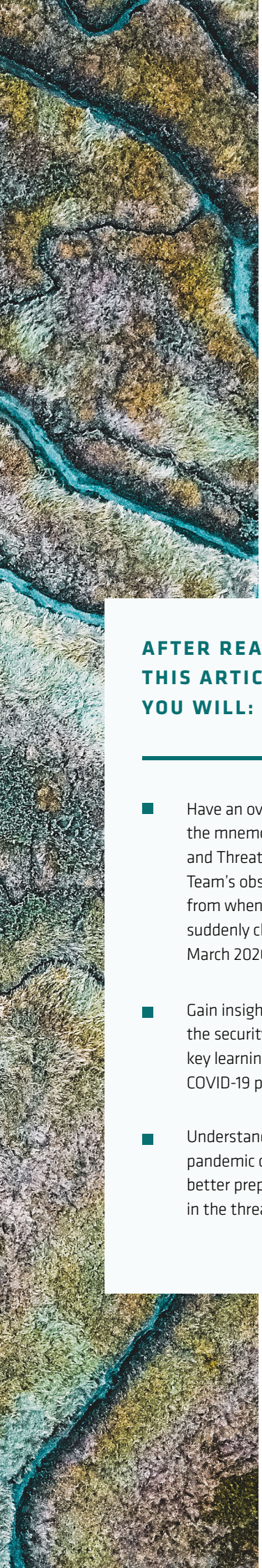
What gives you hope for the future of security?
We are seeing a lot of interesting new initiatives that focus on protecting SME all over Europe. Hopefully, the supply chain companies mentioned in the two other questions can get access to the same technology that now is available for banks and larger government organisations, without the full costs it would normally have.

ARTICLE

Lessons learned from COVID-19

A threat intelligence perspective

Marie Elisabeth Gaup Moe, PhD.
Senior Threat Intelligence Consultant



AFTER READING THIS ARTICLE, YOU WILL:

- Have an overview of the mnemonic SOC and Threat Intelligence Team's observations from when the world suddenly changed in March 2020
- Gain insight into some of the security community's key learnings from the COVID-19 pandemic
- Understand how the pandemic can make us better prepared for shifts in the threat landscape

When looking at the cyber threat landscape in 2020, we cannot ignore the impact of the global pandemic. COVID-19 has affected us all, both in our personal and work lives. From a security perspective, COVID-19 is also an interesting case as it lets us study how both people and threat actors behave in times of disruption. Looking back at the start of the pandemic, it's also worthwhile looking into how prepared we were when this disruption forced us to change our work habits, or at least in which areas we passed the test and which areas we didn't.

This article will explore exactly that; how is security affected when faced with a global crisis, and most importantly, how we can take these lessons forward to prepare for a future that will continue to be unpredictable?

As people and societies have kept adapting to the "new normal", so have (unsurprisingly) cybercriminals and threat actors. Let's start by looking at some of the observations we've made regarding how the pandemic affected the threat landscape and overall security posture of organisations and companies.

Working from home impacts the threat landscape

As the lockdowns started in Europe in the spring of 2020, most of the workforce suddenly found themselves working from their home office. This impacted the threat landscape in various ways.

Many organisations were not prepared for a scenario where most of their employees would connect remotely into the company network. Many had to scramble their resources to equip all employees with the necessary tools, and the use of "bring your own devices" and ad-hoc solutions increased the cybersecurity risk and exposure in the form of more vulnerable devices and more attack surfaces.

Organisations that had already taken most of their infrastructure into the cloud might have had an easier time with the transition. However, many ran into issues with the scaling of services and perhaps running out of user licenses. In addition to a larger attack surface, the increase in people working from home also caused changes in visibility and monitoring of systems. Endpoint detection tools and logging of VPN and cloud proxy solutions became more essential tools for visibility than traditional network monitoring. Some companies also struggled with asset management and maintaining regular software updates of devices that suddenly resided mostly outside of their internal network.

Another issue organisations were faced with was the massive and sudden increase in "external" traffic towards internal resources, which meant it was easier for an attacker to blend in. In short, it became more challenging to discover abnormal traffic since it was all abnormal. ▸

Remote login under attack

Remote login to email accounts, company chat, video conferencing tools, VPN solutions, and other remote access solutions quickly became desirable targets for attackers. Multi-factor authentication (MFA), strong password policies, secure configuration, and patching of all Internet exposed servers running services such as VPN, Remote Desktop Protocol (RDP), or SSH became crucial. We have observed an increase in scanning for vulnerable systems where MFA could be bypassed, coupled with "password-spraying" or "credential stuffing" where attackers systematically attempt to break into systems using popular or previously leaked passwords.

In addition to the password guessing activity, we also saw a number of vulnerabilities being exploited in popular network perimeter solutions such as Citrix NetScaler, F5 BIG-IP, Pulse Secure VPN, and Fortinet VPN.

Users susceptible to COVID-19 phishing

The added stress of dealing with the pandemic combined with working from home disrupting our usual work-life separation might also have affected our security posture, making us more likely to fall prey to the attackers' lures. Cybercrime actors capitalise on the sense of urgency and fear. This is why pandemic themed clickbait seems particularly useful, especially when disguised as company internal, official, or government communication.

Starting in March, we saw cybercrime threat actors increasingly adding COVID-19 themes to their phishing campaigns, SMS phishing, business email compromise (BEC) fraud, and ransomware. We did not observe an increase in the absolute number of phishing campaigns. However, we did observe some threat actors using existing infrastructure and repurposing their malware to include pandemic lures.

Ransomware is increasingly problematic

The healthcare sector has been under tremendous pressure. To add to this, some cyber criminals decided that hospitals and medical supply chain companies were perfect targets for ransomware attacks. Probably because of an increased willingness to pay a ransom due to the importance of keeping critical systems available at all times. In 2020, we sadly saw the first reported death of a patient due to ransomware. The incident shut down critical systems at a German hospital, leading to a patient needing urgent care being redirected to a different hospital and not getting there in time.

Ransomware has become increasingly problematic, and we have observed how criminals have streamlined and adapted their operations to become even more effective in their moneymaking schemes. Some are even extorting their victims more than once by also exfiltrating data and threatening to leak sensitive company data if the victim does not pay the ransom.

The use of so-called "initial access brokers" has also become increasingly popular. These brokers do the groundwork by phishing or scanning for vulnerabilities, harvesting credentials or installing malware, and gaining a foothold with the victims, where they later sell the access to their victims to ransomware "as a service" operators. There are several malware infections that have been found to be precursors of ransomware. Operators of malware botnets like Emotet and Trickbot seem to have partnered with ransomware gangs such as the operators behind the Ryuk ransomware.

Something for security professionals to look out for is unauthorised installations of the pen-testing tool Cobalt Strike, which is one of the favorite tools used by criminals for controlling multiple infected endpoints within a network before launching devastating ransomware attacks. Early detection and removal of the initial access intrusions, together with a sound and tested backup strategy, seems to be the best way to be prepared for handling these incidents.

COVID-19 targeted attacks

Nation-state threat actors also started utilising pandemic themed lures in their usual targeted attack campaigns. Some threat groups seemed to have been explicitly tasked to carry out espionage campaigns targeting COVID-19 crisis response. The race for a cure or a vaccine created an obvious intelligence gathering need, and targets that have been publicly reported include medical research facilities, the World Health Organisation, and government emergency response agencies. Such targeted attacks can not only provide adversaries with research data to help them speed up their own development initiatives but more concerning, they can be used to attack the integrity of the vaccine research programs and the vaccine supply chain by modifying, not exfiltrating, data potentially delaying or halting projects.

Defenders joining forces

The landscape described so far is gloomy, but we should remember that a crisis is also an opportunity to change things for the better. In response to the rise of cybercriminals taking advantage of the situation, defenders started joining forces. Fighting back with information sharing as our weapon.

Two good examples of which are the Cyber Threat Intelligence League (CTI)[1] and the COVID-19 Cyber Threat Coalition[2]. They were formed as volunteer groups acting as information sharing hubs for indicators of compromise, takedowns, triage, and law enforcement escalations of COVID-19 related cybercrimes.

Lessons learned

Even with the most common security enhancements introduced by many organisations this year, remote workers are still not protected by the same level of security as they are

[1, 2] *See Reference List at the end of the report*

when working from their office. They are, for example, more vulnerable to phishing attacks as there are no colleagues around who can help validate an email with suspicious URLs or attachments.

By the end of 2020, it seemed that the majority of organisations had resolved their scalability issues and the practical hurdles related to a remote workforce. Reviewing and optimising security controls is now a highly recommended next step to ensure users stay safe and that the attack surface is minimised.

While patch management, strict security policies, robust authentication, and proven VPN solutions are mandatory mechanisms in a remote workforce scenario, the need to keep delivering internal security awareness training is critical and should not be underestimated.

Visibility into what is happening on the endpoints is also critical. Ensuring that all core applications are not only logging details about usage but also that these logs are consumed by a centralised log management solution is very important.

Outsourcing contracts should also take into consideration agreements with sourcing partners that move their operations from customer approved office locations to their staff's home offices. This has been an eye opener for many, and in addition, GDPR further complicates outsourcing for many organisastuoins, as the Schrems 2 ruling renders the Privacy Shield agreements concerning the transfer of personal data between the US-EU/EEA invalid. If that wasn't enough, uncertainties related to Brexit further complicate outsourcing of datacentres and IT operations.

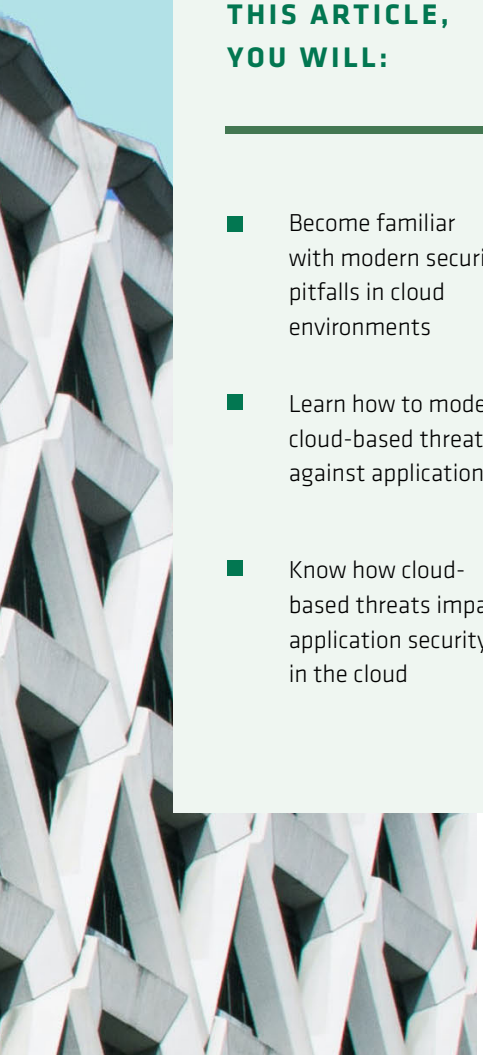
While we are hopeful that this pandemic will have a marked end date for the history books, threat actors will persevere, continuing to adapt and take advantage of current events and the worries that are on the minds of people and organisations. Thus, we will need to continue to prepare for the "unpreparable" in order to fight back against cybercriminals and other threat actors. The next time around, however, we do this strengthened by the bonds formed and the lessons learned during this time of crisis. ●

ARTICLE

Cloud is not just somebody else's computer

New paradigms for security threats in modern cloud applications

Cody Burkard
Senior Cloud Security Consultant



AFTER READING THIS ARTICLE, YOU WILL:

- Become familiar with modern security pitfalls in cloud environments
- Learn how to model cloud-based threats against applications
- Know how cloud-based threats impact application security in the cloud

Make no mistake; application development for public cloud infrastructure is the new norm. Whether this is because of the speed of development and the lack of infrastructure maintenance, the native automation capabilities in cloud environments, or a variety of other factors, it is safe to say that application development in the cloud is here to stay. This leads to the question: what new security considerations are there for cloud-native applications?

Modern cloud security issues

The term "cloud security test" may invoke a number of thoughts and assumptions depending on whom you speak with. To some, this simply means web application testing when the application is hosted in a cloud environment. To others, this could mean a traditional penetration test, where the goal is to gain access to a certain objective within infrastructure on a Virtual Network. We won't debate the details here, but for the sake of this article, we choose to define a cloud security test according to the following goal:

"Identify security weaknesses and vulnerabilities in the cloud environment that would allow an adversary to impact the confidentiality, integrity, or availability of any service within or otherwise dependent on the environment"

This definition is intentionally quite broad. It is almost as vague as the term "cloud security test" - it encompasses almost any type of vulnerability that could exist within the cloud or its infrastructure, be it on a virtual machine or through a cloud configuration. However, the results of these tests commonly centre on a few security issues. Specifically, amongst the most prevalent issues in cloud deployments are networking misconfigurations, poor secret management, and Identity and Access Management (IAM) misconfigurations. Let's take a closer look at why each is prevalent in cloud environments.

Network misconfigurations

Most cloud resources in a modern cloud provider can be connected to a network, and it is possible to enforce network access control lists (ACLs) on inbound or outbound connections to the resource. However, it is also possible to run most resources without connecting them to a virtual network. Many times, this is the default behaviour for cloud services.

For example, Storage Accounts in Microsoft Azure and S3 Buckets in AWS are by default accessible via direct network connections from the Internet. While both require authentication in order to connect, an adversary with a stolen secret could connect from their own environment. To limit network access, custom bucket policies in S3 or IP address whitelisting in Azure Storage must be utilised to limit network connectivity. During cloud security reviews, we frequently observe that developers believe direct internet connections to cloud resources, such as storage, are safe. Common reasoning behind this assertion ▸

The focus on internal threats is not a novel concept, and is very much in-line with traditional approaches to offensive security on technologies such as Active Directory. However, the nuances in the details are interesting.

follows that the cloud provider manages the storage services, and because storage services still require authentication to access the data. At first glance, this logic seems reasonable because Microsoft will automatically patch vulnerabilities on the host, and the secrets for the storage services are cryptographically secure, so traditional attacks seem infeasible. However, as the rest of this article shows, there are additional security concerns in cloud infrastructure that cloud administrators and developers should consider.

Poor secret management

Most cloud resources require a secret in order to connect or interact. In Azure, for example, authentication with a secret key is required on storage services, caching services, messaging services, and managed API endpoints in cloud environments, as well as many other frequently used services. As an application developer in cloud environments, this means that these secrets need to be stored somewhere safe, automatically updated, and should never be leaked.

Because of the abundance of secrets and the constant need to update them, managing and protecting these secrets is a difficult task in cloud. During development and deployment, it is very common for plaintext secrets to be stored in storage services, local developer systems, or in configurations throughout the environment. Keep in mind that when an adversary steals one of these secrets, they can authenticate to the resource to which it belongs.

IAM misconfigurations

One of the most significant differences between cloud infrastructure and on-premise infrastructure is the management model for each. In the cloud, the traditional, physical installation process of new hardware has been replaced by web APIs. In essence, each cloud is one huge web application that integrates with sophisticated virtualisation infrastructure as a backend, and lets users create their virtual environments via a web interface or API.

This means that user permissions on this "web app" are extremely important. A developer should not have access to modify the network, in the same way that developers in on-premise applications do not have access to the datacentre. In practice, each cloud provider governs user permissions differently to solve this problem. For example, AWS uses IAM policies while Azure uses role-based access control (RBAC) roles. However, the principle is the same: limit user access to specific cloud services, and to specific actions on those services.

However, as cloud environments grow, these IAM assignments become more complicated. Maintaining fine-grained IAM assignments creates management overhead, and frequently leads to mistakes. If an adversary already has limited access to an environment through a previous vulnerability or as an insider, IAM misconfigurations can lead to the adversary gaining more access to the network than they would with a proper configuration. In some cases, misconfigurations may even allow an adversary to escalate their IAM privileges.

This is a particularly dangerous type of misconfiguration, but its impact is also very dependent on the IAM role that an attacker can access. For example, it is more dangerous if an attacker can add new data to a storage service than if an attacker can simply read cloud configurations. We will explore this further in a later section.

Identity is the new perimeter

A brief analysis of the three security issues discussed above yields an interesting characteristic about cloud security: most security issues in cloud focus on defence in depth. Whether we discuss network attack surface, secret management, or IAM, the insider threat or an adversary with a compromised user account is the most likely to exploit each. This supports one of the modern security principles in the cloud: identity is the new perimeter.

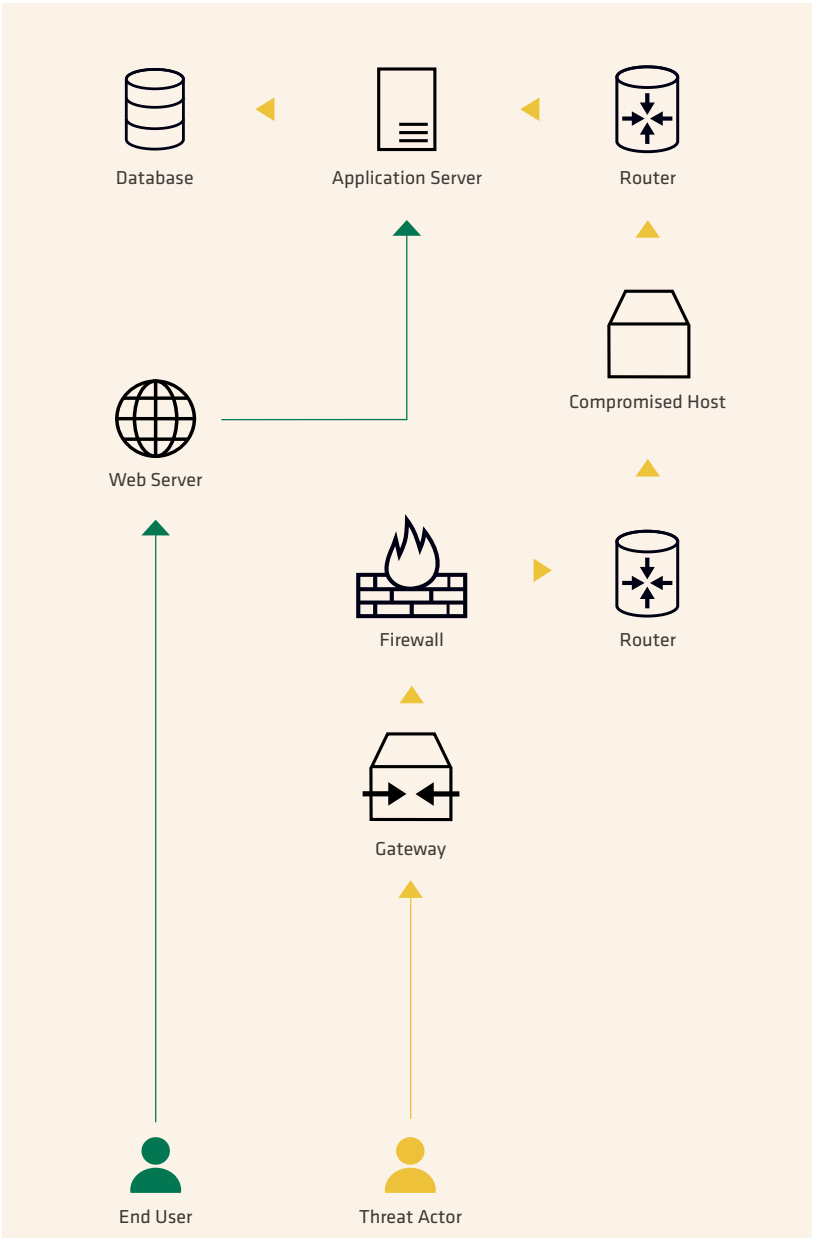
This assertion of identity being the new perimeter is also supported by the prevalence of attacks we see – namely phishing attacks, which are often targeted specifically towards capturing identities. Most threat reports highlight phishing as one of the main techniques employed by sophisticated threat actors, likely due to the high success rate. If the impact of a phishing attack is a stolen IAM role in

cloud, and the attacker has access to modify infrastructure via API calls, it is no wonder that securing user identity has become such a large focus.

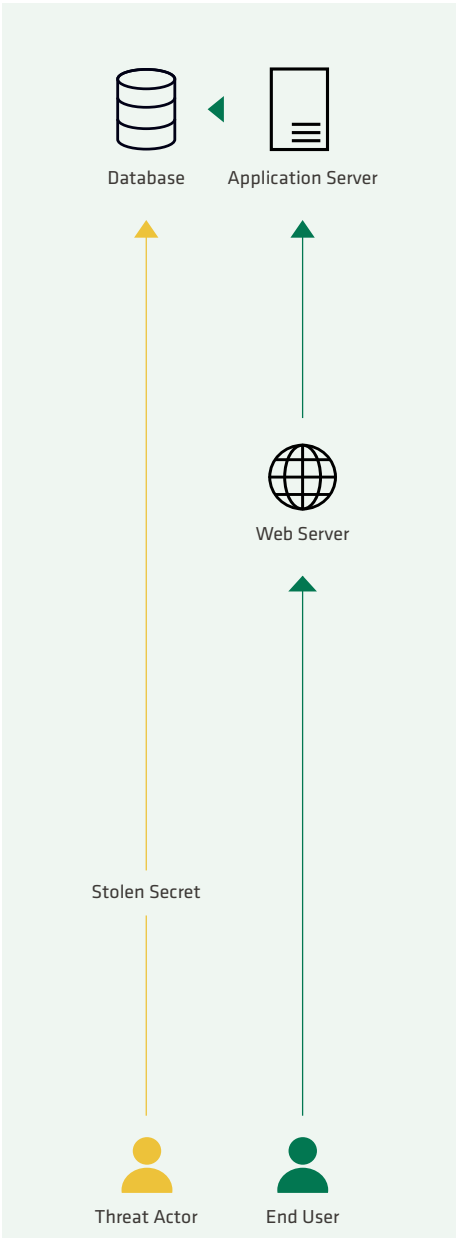
The focus on internal threats is not a novel concept, and is very much in-line with traditional approaches to offensive security on technologies such as Active Directory. However, the nuances in the details are interesting. For example, consider a scenario with both excessive network exposure and an IAM based attacker. If the attacker escalates their privileges entirely within the cloud provider IAM roles, they can directly connect to internal resources without the need to pivot or move laterally through the network. This situation makes the attacker's job much easier by removing all of the obstacles in the network, and making it easy to maintain persistent access to the internal resource without compromising any hosts.

The diagram below shows two example scenarios that highlight this point. In scenario B, the IAM-based attacker can directly access a cloud database with a stolen secret. In scenario A, the traditional attacker must traverse a number of network devices to gain access to the database.

If this type of thinking still seems outlandish, consider a real world case study: the breach of Capital One in 2019 that resulted in the loss of personal information for over 100 million customers. A public FBI complaint includes some details from a technical analysis of the breach, which is enough to gain ▸



Scenario A: Example of an attacker gaining access to on-premise application database



Scenario B: Example of an attacker gaining access to cloud application database

a basic understanding of the attacker's path. To begin with, a small diagram of one potential representation of the Capital One environment is included below in Figure A.

The S3 Bucket is not connected to a virtual network. The WAF is running on an EC2 instance (Virtual Machine) in AWS, and has an IAM role attached, as seen in the diagram. The IAM role allows all actions on all S3 Buckets in the account, since it's using wildcards, which means it can access all data on those S3 Buckets. So what does this mean in practice?

The scope and permissions of the WAF's IAM role make it possible for the WAF to read the contents of the internal S3 Bucket, while the public network access of the S3 Bucket makes it possible for an authenticated user to read the contents of the bucket directly over the internet. This creates a dangerous scenario, where an attacker with access to the WAF can impersonate the WAF's identity, and access the S3 Bucket from their own location. The industry broadly assumes that this is exactly what happened. Most analyses suggest that a server-side request forgery (SSRF) vulnerability allowed the attacker to steal an IAM temporary credential from the WAF, and directly dump the contents of the S3 Bucket, as seen in Figure B.

Based on the presumed setup in AWS, the root cause for this issue lies in the overly permissive IAM role on the WAF and the public network availability of the S3 Bucket. The SSRF in the WAF gave the attacker the first step into the environment, but the misconfigured network and IAM policies made it possible to pivot to new resources. In short, this breach offers a simple lesson that we as an industry have known for years: traditional security principles such as defence in depth always apply, and we simply need to adopt them to new environments.

Because of the centrality of IAM within cloud security, attackers are more focused on acquiring secrets and credentials, such as the temporary access token, or escalating their privileges to access more capabilities on the cloud APIs. By escalating privileges, adversaries can avoid the need for complex application-level vulnerabilities that are only

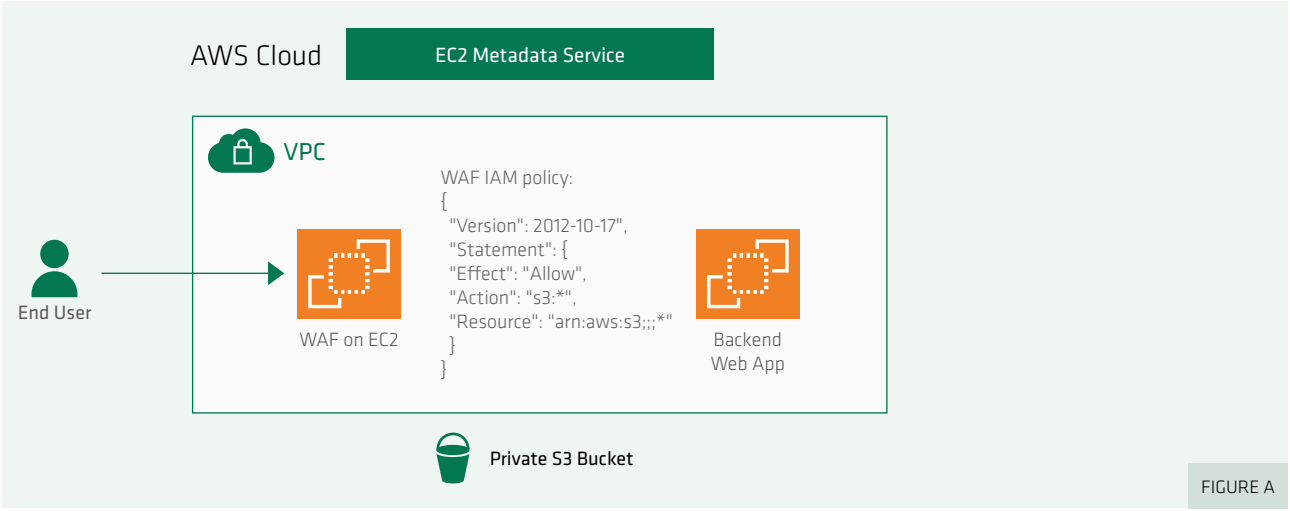


FIGURE A

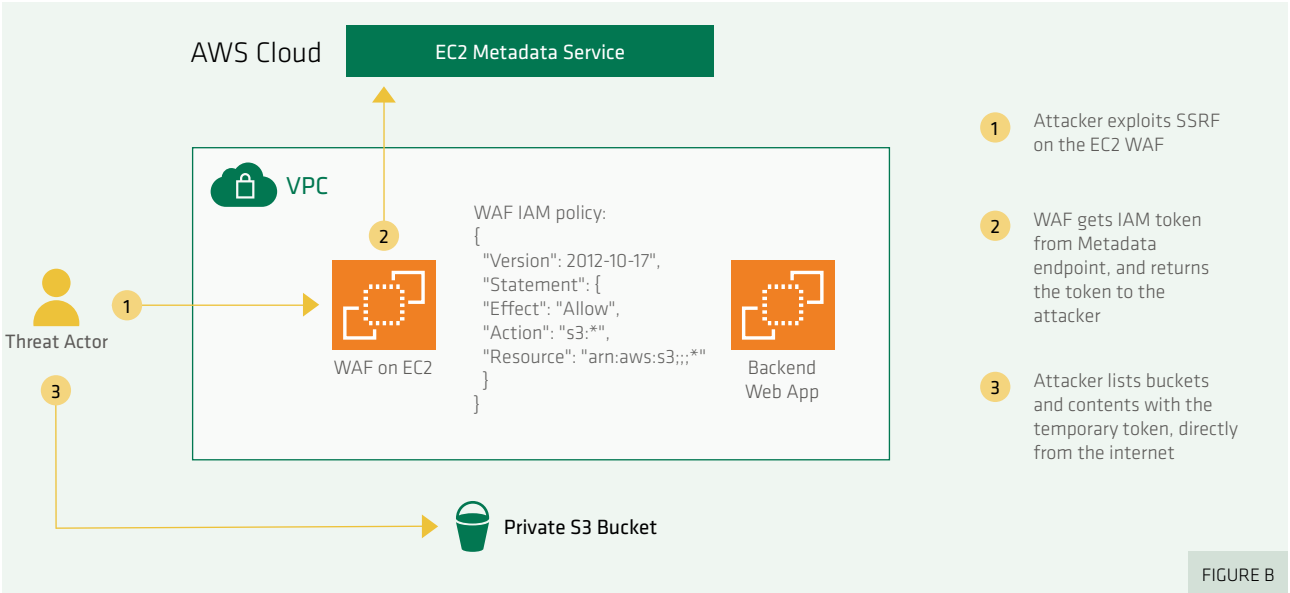


FIGURE B

exploitable at the network layer. This also provides them with new attack capabilities, such as directly modifying existing infrastructure in the cloud.

The impact on application security in the cloud

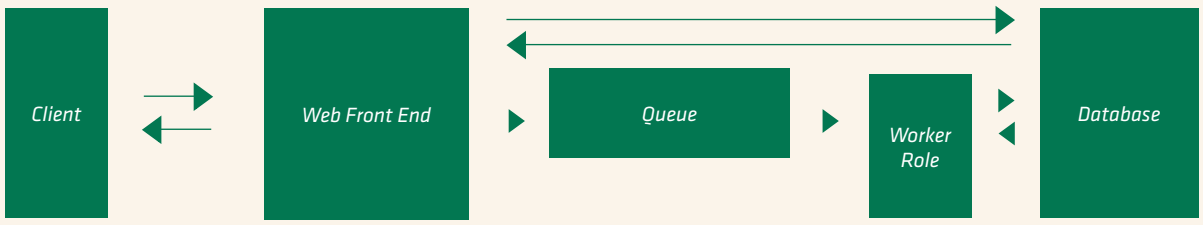
So how should this influence the mindset of application developers? We are talking about IAM and networking, but how much of this is the developer actually responsible for? And how do these issues impact their applications? The short answer again comes back to defence in depth and essentially reducing the blast radius. Developers should at a minimum be aware of these threats, and ideally assume a breach scenario to limit the impact of an internal threat. Let's consider a practical example.

Consider an API that handles banking transactions, which is built in a WQW architecture within Azure. This API should never suffer downtime, and must provide non-repudiation on all transactions. This means that it must be possible to validate the authenticity of each transaction and verify that each transaction occurred as was intended. The consequences for attacks against such an application are high, and we can assume the presence of a sophisticated threat actor.

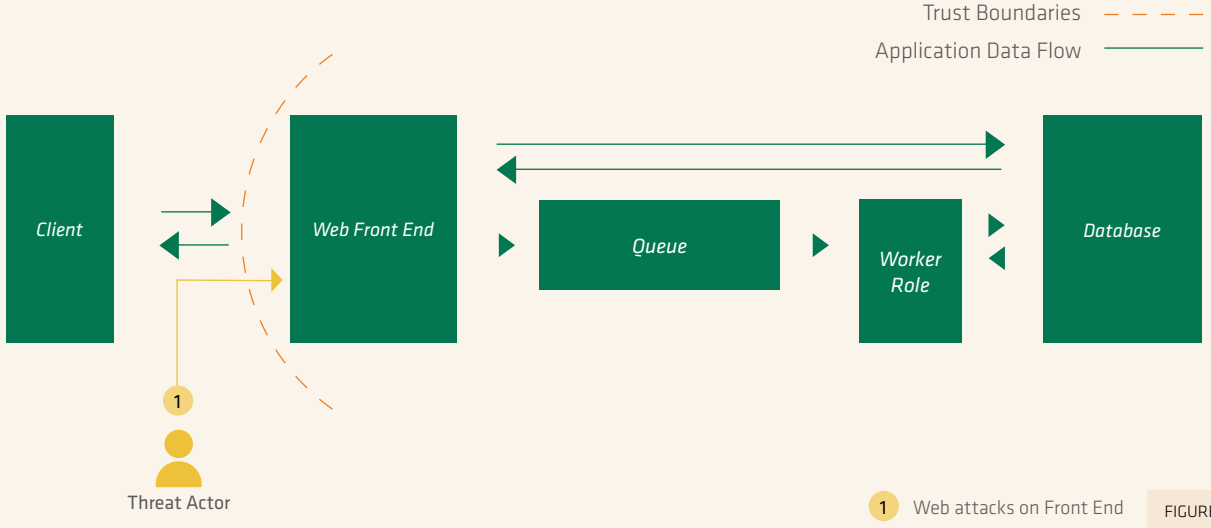
A traditional approach to the threat modelling of such an application may look like what you find in Figure C.

Based on this diagram, we assert that the worker role only polls from the queue, and the queue only contains data from the web-worker. While attacks such as second-order ▸

What is Web-Queue-Worker Architecture?



A WQW architecture is built to take advantage of scalability and managed messaging services in cloud environments. These architectures include three main components: a front-end, a queue, and a worker role. The lightweight HTTP-based front-end service normally handles requests with low-computational requirements, such as pushing new data to the internal queue, or communicating with an external microservice. The worker role handles complex business logic and long-running computation tasks. These two services can easily be scaled up or down using cloud-native features. Most public cloud vendors provide some type of messaging system as a PaaS offering. Often there are multiple options. In Azure, you may use Storage Queues, Service Bus, or Event Hubs. AWS includes options such as Simple Queue Service (SQS), Simple Notification Service (SNS), or Kinesis. GCP offers Task Queues and Pub/Sub.



1 Web attacks on Front End

FIGURE C

injection against the worker role are possible, we can just prevent them with sanitisation and validation on the web role, which protects the worker role. This could be a reasonable approach in a traditional on-premise application security architecture.

However, consider the cloud-based threat actors that possess a stolen secret or have escalated to privileged Azure RBAC roles. Those IAM roles may provide an attacker with new attack surfaces on the application. If the IAM role allows the attacker to modify the network, the adversary may even whitelist their own IP on the queue, giving themselves direct network access. This is an entirely new threat scenario, based entirely on a user with access to Azure roles.

Attacking from the inside out

Based on a threat actor in this scenario, we must adjust the previous threat model. However, the new threat model is entirely dependent on the attacker's capabilities, which in turn are dependent on the RBAC roles or secrets they acquired.

So, let us consider what an attacker's capabilities would be if they acquired some of the built-in Azure roles. Figure D shows that if an attacker gains access to either the Owner or Contributor roles, they can make administrative changes to the entire application, so that developers cannot add any protection. Attackers with lower privileges that are limited to accessing the queue, such as Storage Contributor, Storage Queue Data Contributor, or Storage Queue Data Message Processor still pose a risk to the application; however, developers are in a position to mitigate these risks.

So let's consider an adjusted threat model that takes these roles into account, and assumes the cloud-based attacker has gained read and write permissions on the storage account using the Storage Contributor role as shown in Figure E.

There are now two new trust boundaries – one on the input of the queue, and the other on the output. This is a much more powerful attacker, because they are no longer limited to external attacks on the web front end. In addition, the

Role	Capabilities	Scope	Developer can mitigate if compromised
Owner	Assign roles, full administration access on all resources	Web role, queue, worker role	✗
Contributor	Full administration access on all resources	Web role, queue, worker role	✗
Storage Contributor	Modify queue network access, list keys for queue	Web role, queue, worker role	✓
Storage Queue Data Contributor	Read/write/delete items on queue	Web role, queue, worker role	✓
Storage Queue Data Message Processor	Read/delete items on queue	Web role, queue, worker role	✓

FIGURE D

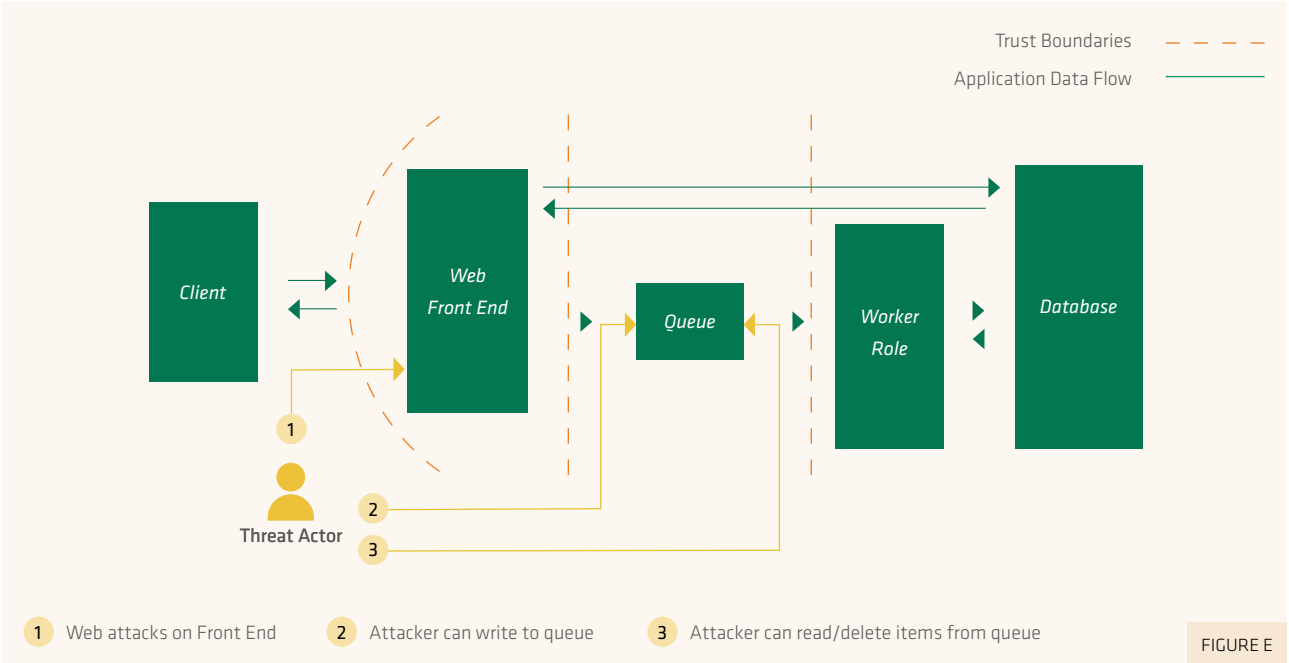


FIGURE E

attacker can read or write data on the internal messaging system for the application. This poses a challenge for the developers of the application: how do you protect against this threat, and what attacks can still be mitigated?

To begin with, consider a few possible attacks in this scenario:
1. The adversary pushes data in an unexpected format to the queue, and attempts to trigger web app attacks on the worker role.
2. The adversary pushes properly formatted data to the queue, and attempts to trigger fraudulent financial transactions.
3. The adversary reads and deletes financial transactions from the queue, preventing them from executing.

Each of these scenarios is catastrophic for the application, and developers should consider how they may protect against them.

There are a variety of ways to protect against these threats, so we won't explore all of the possibilities. A simple approach to mitigation may resemble the following:
Threat 1 and 2: Sign data using the web role, and verify the signature of each transaction on the worker role. Any time the worker cannot verify the signature, it should not process the transaction, and it should trigger an alert.
Threat 1 and 2: In the worker role, perform validation on the data format and validate all data types received from the queue.
Threat 1: Verify the signature before performing any dangerous function in the worker, such as deserialisation.
Threat 3: Number each transaction in the web role, and sign the transaction number along with the transaction data. The worker role should verify that no transaction in the chain is missing.

The takeaway here is the attack surface on an application changes when considering IAM-based threats in the cloud, and application developers are in a position to, and should consider how to protect against these types of threats.

Key takeaways for offense and defence

Secure application development in the cloud has fundamental differences with the traditional approach to application security. The primary difference is the IAM-based and stolen secret-based adversaries, as well as the prevalence of network exposure issues in the cloud. This means that the attack surface of an application is constantly in flux, and depends on each modification to IAM roles within the cloud environment. Within the security profession, employees in each role need to consider these differences and understand how they should adapt their approach to security in order to solve these challenges.

For Developers:

Developers should use threat models to highlight these differences during the development process. During this threat modelling process, consider the following questions:

- How could a stolen secret affect this application?
- How could IAM-based attackers with different permissions attack this application?
- Can network access to this application be limited further?
- How can defence in depth principles be applied to the application development process to protect against cloud-based threats?

For Security Architects:

Security Architects face the challenge of managing IAM roles in their organisation. It is important for architects to adopt centrally managed IAM and enforce the principle of least privilege to prevent these issues from arising. This becomes particularly challenging as more organisations adopt multi-cloud approaches, so consider cloud-agnostic products for monitoring, enforcing, and automating the management of IAM roles.

For Penetration Testers:

Penetration Testers and offensive specialists should consider the cloud attack vector while assessing these applications. In our WQW example, a pentester could easily miss a high-severity vulnerability that is only exploitable with an IAM role in the cloud. This is a good argument for requesting access to cloud-based users with IAM roles to help perform a web application assessment.

During offensive engagements, always consider the following:

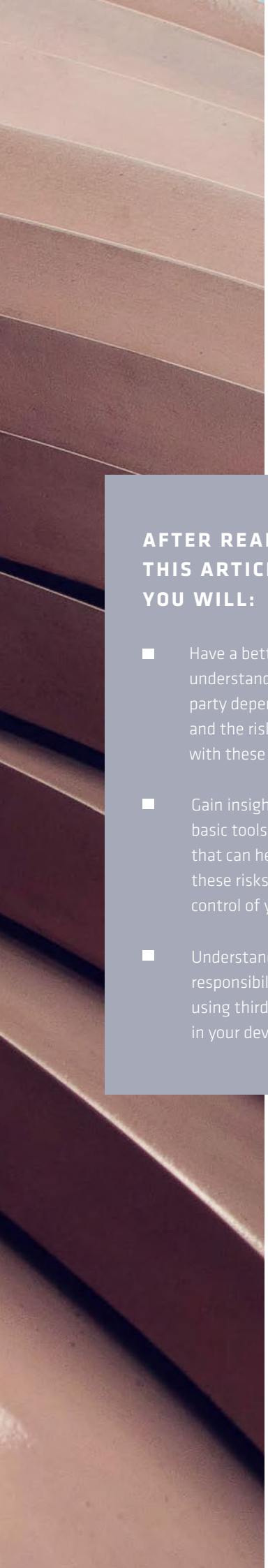
- What is the architecture of this application?
- What are all the ways I can interact with the internal components of this application?
- Have I compromised any users with IAM permissions in the cloud?
- Can I acquire new IAM permissions using any exploits I have identified?
- How can I find or extract secrets to internal components of the application?

Overall, for both offense and defence, consider that cloud does not change the fundamental principles of security. The concepts are all the same, but application of these concepts is different in cloud, and so are the consequences of neglecting them. To conclude: it doesn't matter if cloud is someone else's computer – it's your responsibility. ●

ARTICLE

Securing third-party dependencies in development

Andreas Claesson
Senior Technical Security Consultant



AFTER READING THIS ARTICLE, YOU WILL:

- Have a better understanding of third-party dependencies and the risks associated with these
- Gain insight into some basic tools and actions that can help navigate these risks and gain control of your code
- Understand what responsibilities follow using third-party code in your development

More often than not, development in large projects means using third-party libraries. According to industry estimates[1], this applies to as much as 85% of the code in a typical application. The complexity of modern projects requires developers to use libraries that are convenient and prevent them from having to reinventing the wheel. However, there are some considerations from a security point of view that need to be taken into account.

What makes securing third-party dependencies so complicated?
Just like with any code, external libraries may contain vulnerabilities. Vulnerabilities in third-party libraries commonly originate from two main categories:
1. *Bugs intentionally introduced by someone with malicious intent*
These are hard to spot "backdoors" in the code that are easy to exploit for those who know about them.
2. *Bugs mistakenly introduced by developers*
Most often however, vulnerabilities in third-party libraries are there by accident.

An interesting example from the first category is the attack on the EventStream library in 2018[2]. This highly popular JavaScript library was compromised, and a third-party dependency was added containing encrypted malicious code. If EventStream was used with a specific cryptocurrency-related library, the malicious code would try to steal bitcoins from your cryptocurrency wallets.

A well-known example of the second category is the Heartbleed OpenSSL vulnerability that was discovered in 2014[3]. OpenSSL is an open-source library that has its own implementation of TLS/SSL, and was, and still is, widely used in web servers, operating systems and hardware appliances. Half a million well-known and trusted websites, like Yahoo.com, were vulnerable because of this bug. In this case, the implementation was flawed, not the TLS/SSL standard itself.

While fixing software bugs might seem like a simple update, in practice it's not always that easy. Should a bug be fixed incorrectly, it can actually increase your overall technical risk. For the library developers, there is a cascading effect as well, since a lot of projects depend on the same open-source library, and even small changes in the library can break other projects' code.

Furthermore, the vulnerable library can also depend on code from other libraries or sources that might be even harder (or impossible due to legacy reasons) to fix. This means you can be affected by dependencies beyond your own third-party dependencies (sometimes called indirect dependencies). To complicate things even further, these libraries can have a different licensing model compared to your own code. This means that there can be restrictions on the use of the code, for example free for private use but not for commercial, or other requirements that need ▸

[1, 2, 3] *See Reference List at the end of the report*

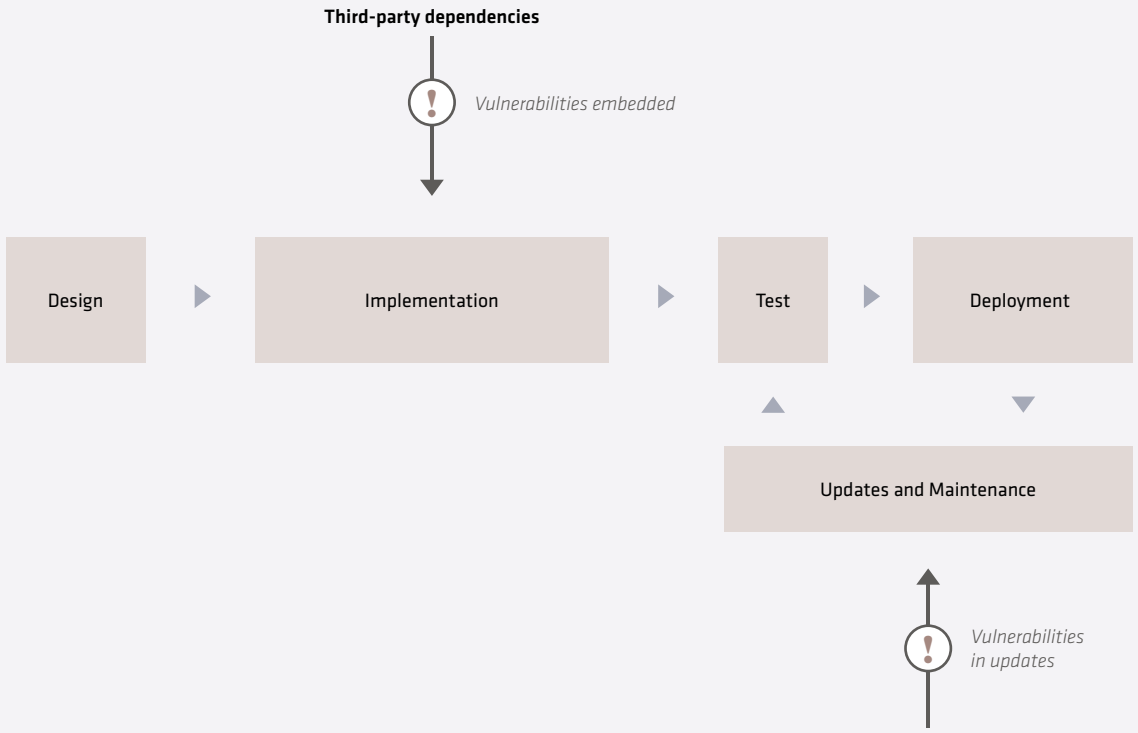


Figure 1: *Most common ways vulnerabilities are being introduced to code*

to be followed. Other models allow free use, but restrict modification of the code, and certain copyright statements might need to be included.

Getting a good overview of the code that is *actually* included in a build quickly becomes complicated. At the end of the day, it's the product owner who is responsible for all the code in a project, regardless of whether it's developed in-house, or from third-party libraries. Below I will explore a few actions and products that can make things easier when navigating this area and getting control of the actual code used.

Be aware of suspicious third-parties

There are literally hundreds of thousands of potential libraries to choose from, and they all present various risk levels. Blindly using the first library that looks promising without a little bit of investigation is rarely a good idea. Some libraries have severe vulnerabilities by accident, some on purpose. Some libraries are updated and maintained often, some have been unchanged for years, and so on. As a general idea, using libraries with a big community and that are well maintained is a good start.

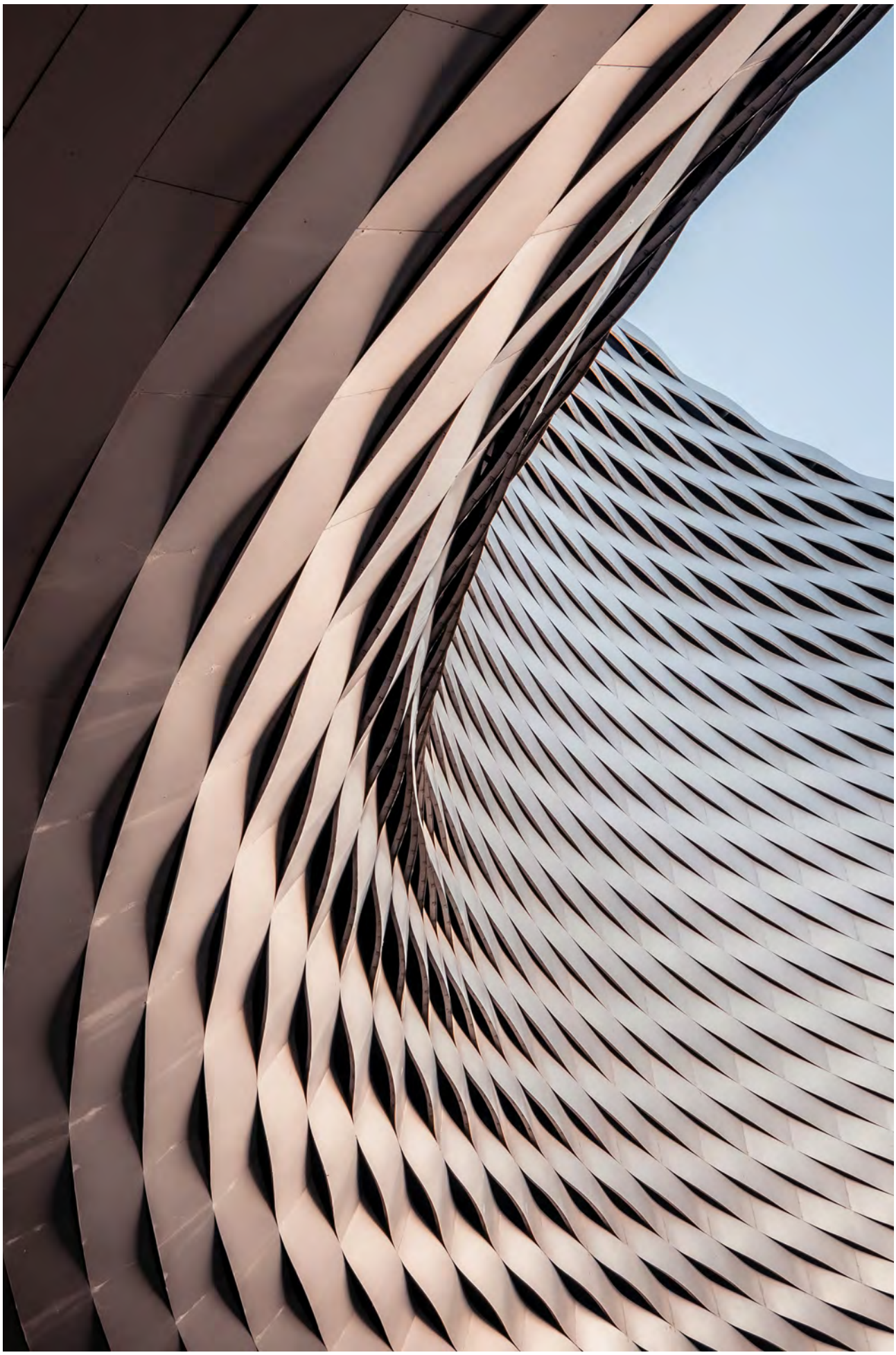
There have been several security incidents in the past where malicious code has been introduced to innocent looking libraries, and later pulled into projects. As Figure 1 shows, the two most common development phases where vulnerabilities are being introduced, both maliciously and by accident, are either during implementation or later on during updates and maintenance.

Using a local repository that doesn't blindly update dependencies remotely is a good option to mitigate this risk. Ensure the local repository uses well-known, stable releases that are not tampered with by getting the code from the original source. If possible, confirm the code integrity with checksums if they are available.

Knowing what version of a specific library you are using is also important for keeping the risk low. Do not be tempted to simply use the latest version out of convenience. Update versions deliberately, and read the release notes carefully before upgrading any libraries. Pining dependency versions in your code ensures they are not updated automatically, and will help prevent malicious updates and backdoors from sneaking in unnoticed. Keeping an inventory of the code used in production is the first step in protecting it. Keeping the inventory up-to-date is even more important!

Automate your build

As mentioned, an important step in securing your third-party dependencies is to know what code you are actually running in production. In order to keep track of vulnerable versions of libraries, you have to know exactly which versions are included in your code. In order to do that, an automated build process should definitely be in place. ▸



While code quality is not necessarily a security problem, writing bad and confusing code can definitely lead to bugs, which in turn can very easily become a security issue.

This will make sure builds are consistent, and will make it easy to include security checkpoints. However, implementing a build process in practice can be difficult, especially if the organisation is not used to working this way. Modern industry practices such as DevOps provide a common solution to this problem.

DevOps automation is becoming increasingly popular. In its simplest form, DevOps is a way to structure build automation. It's often described as a "pipeline", where source code enters in one end, and a production-ready build comes out the other. Most often the source code comes from a repository with version control like GitHub or similar. The pipeline then consists of various "stages" like code build, test, and so on. This drastically shortens the time from development to production, and is one of the main reasons why it's becoming more popular, and arguably a standard practice today. A DevOps pipeline helps keep track of where code comes from, which is essential for keeping it secure.

Having a well-designed DevOps pipeline is key here, and that includes everything from source control to build scripts and artifact repositories. This will also provide a solid foundation for introducing additional security measures. For example, the code base can be scanned for vulnerabilities when committing new code, and the final artifact can also be scanned for vulnerabilities when the build is finished. This will make it easier to keep track of the code running in production, and potential risks with it.

Scanning your dependencies

As described earlier, third-party dependencies can quickly become complicated. The issue grows exponentially when you take into consideration dependencies of dependencies (indirect dependencies). In some cases, vulnerable libraries pulled in through other dependencies might not be directly addressable other than by communicating with the upstream vendor. Simply fixing the problem in the indirect dependency might break the upstream vendor's code, so a seemingly simple bug can sometimes be very problematic to fix. However, to fully understand your security posture, it's important to be aware of all vulnerabilities found in your dependency chains.

For a medium to large size project, trying to detect and track vulnerabilities throughout these dependency chains manually is impossible. Mapping out all libraries and the connections between them quickly becomes overwhelming, and it's easy to lose track of what goes where. Luckily, there are solutions that can help, both commercial and open-source. Since most of the tools provide false positives to some extent, time and effort for initial tuning should be expected. However, that is a one-time effort that's often worth the initial investment. ▸

Throughout the DevOps pipeline there are plenty of opportunities to implement these tools as security checkpoints, but exactly where they should be introduced depends heavily on the end product and the rest of the development chain. For example, the code base (and its dependencies) can be scanned for vulnerabilities at every commit to the code repository.

Starting out with a simple open-source dependency checker such as OWASP Dependency Check[4], can help lower the risk significantly with rather minimal effort. This is a tool that simply checks the dependency tree file for the versions that are being used in the code, and compares that information to a list of vulnerabilities present in different library versions. Another great open-source alternative is RetireJS[5]. Both of these work in a similar way, and can easily be integrated in a DevOps pipeline.

A possible next step in the process may be to look for commercial alternatives, such as Snyk or BlackDuck. Commercial products usually have their own database of vulnerabilities that may provide more accurate results. In addition, when new vulnerabilities are discovered, these databases tend to be updated immediately with new detection patterns.

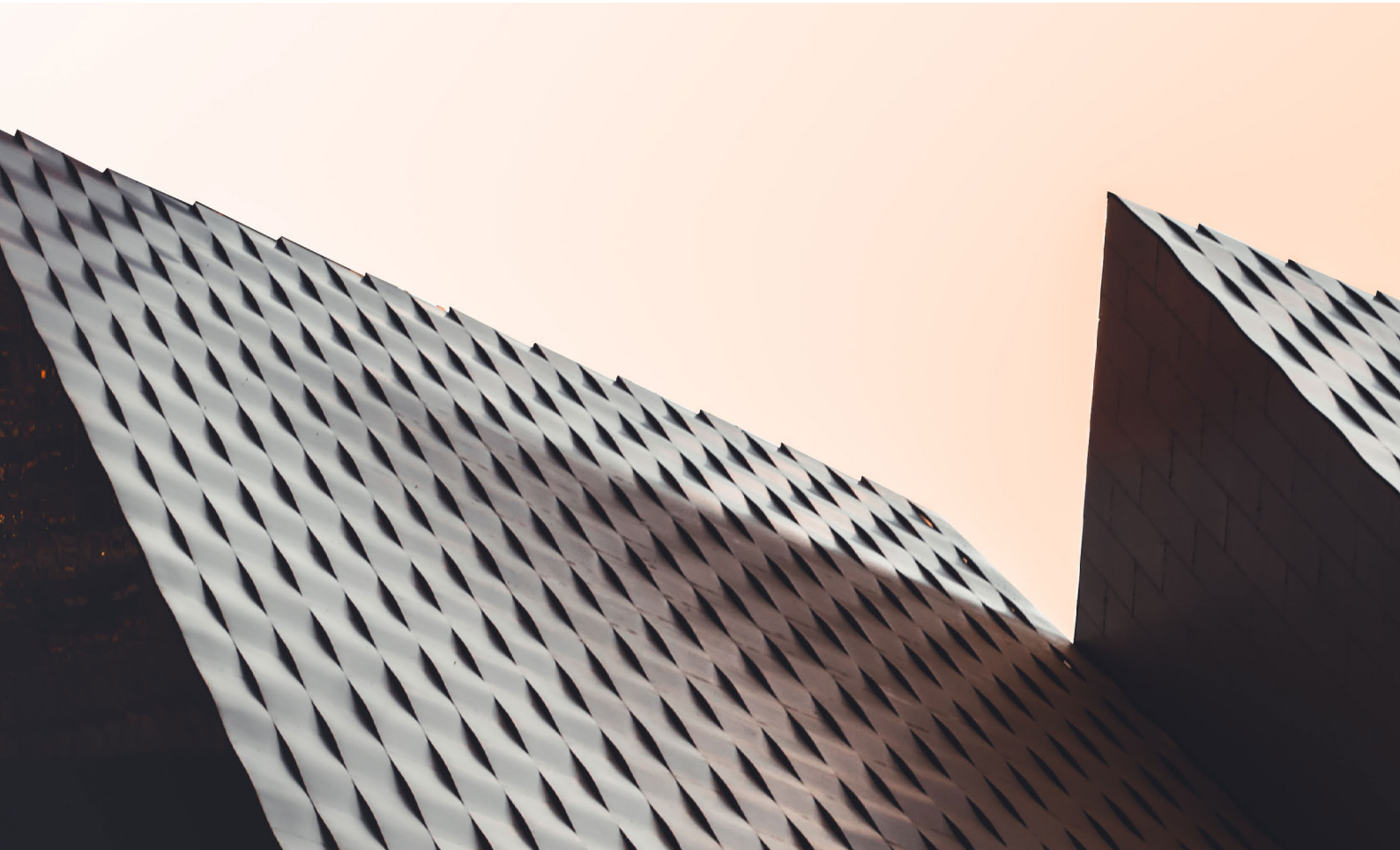
An additional measure that can be taken during the development process is using so-called "linters"[6] to continuously check for basic code quality issues such as bugs, syntactic errors and suspicious constructs. It can also be used to highlight code that doesn't conform to a specific code style standard set by a company, in order to make code consistent across all developers in a department. While code quality is not necessarily a security problem, writing bad and confusing code can definitely lead to bugs, which in turn can easily *become* a security issue.

Regular code reviews

Another helpful process one can use to get control of your third-party dependencies is to augment automated dependency scanning by performing a manual review of the code. This can be done both internally and by using an external auditor. Internal code review is more focused on knowing what third-party libraries are being used, and which parts of your code base depend on them. As mentioned previously, knowing what code you run in production is key.

Using an external auditor can help you focus on the security aspect of the code, providing insights and knowledge about how to write secure code and avoid mistakes. An external source code review is a thorough process, and utilises automated tools (e.g. Static Application Security Testing, or SAST) as well as the software security experts' experiences. Source code reviews can be performed periodically, for example once or twice a year, or if the product is in early development, whenever new functionality is deployed.

[4, 5, 6] *See Reference List at the end of the report*



It can be difficult to determine the severity of a vulnerability, and whether it can be exploited, or if it's only affecting a part of the code not currently being used. Sometimes a vulnerability in a dependency cannot be easily fixed. Here a code reviewer would need to investigate the code to determine if the vulnerability really needs to be fixed, or if it's appropriate to make a note, and accept the vulnerability and risk associated with it. In these cases, an external auditor with specific security knowledge can prove helpful.

Summary and recommendations

Securing third-party dependencies is a complex topic, and it can be difficult to know where to start. Based on the discussions above, I recommend considering these three suggestions:

- *Knowing the source*

It's important to know what version of a specific library you are using in order to keep the risk low. It's equally important to make sure you update the libraries deliberately, and not default to using the latest version available when building the code. Use a local repository that doesn't blindly update dependencies remotely.

- *Build automation and DevOps pipeline*

Knowing what code you run in production is key, and automated builds can really help you gain an overview. Having a well-designed DevOps pipeline makes it easy to implement checkpoints in the form of automated scanning with either commercial software or open-source tools.

- *Regular code reviews*

Internal code reviews make sure you know what third-party libraries are being used, and which parts of the code base depend on them. This can also include automatic SAST tools in the DevOps pipeline. Security code reviews using an external auditor provide insight into security problems in the code, and could be done regularly in order to keep code quality high.

As modern development projects grow more complex, the benefits of using third-party libraries are becoming even more evident. This means we need to pay closer attention to the security considerations associated with using these libraries. Hopefully, this article has introduced, or reminded you about, some tools and actions that can help you navigate these risks and gain control of your code.

REFERENCE LIST

Enterprise Security Architecture: Optimise your security investments

1. https://sabsa.org/sabsa-executive-summary
2. https://attack.mitre.org
3. https://www.sabsa.org
4. https://www.nist.gov/cyberframework
5. https://www.mnemonic.no/risk-services/security-strategy/
6. https://www.iso.org/isoiec-27001-information-security.html
7. https://www.sans.org/critical-security-controls
8. https://cloudsecurityalliance.org/research/cloud-controls-matrix/
9. https://nvd.nist.gov/800-53
10. https://nsm.no/regelverk-og-hjelp/rad-og-anbefalinger/grunnprinsipper-for-ikt-sikkerhet-2-0/introduksjon-1/

We need to talk about insider threats

1. https://cmmiinstitute.com/learning/appraisals/levels
2. https://nsm.no/regelverk-og-hjelp/rad-og-anbefalinger/grunnprinsipper-for-personellsikkerhet/introduksjon/

Lessons learned from COVID-19: A threat intelligence perspective

1. https://cti-league.com/
2. https://www.cyberthreatcoalition.org/

Securing third-party dependencies in development

1. https://www.sonatype.com/hubfs/SSC/2019%20SSC/SON_SSSC-Report-2019_jun16-DRAFT.pdf
2. https://www.synopsys.com/blogs/software-security/malicious-dependency-supply-chain
3. https://news.netcraft.com/archives/2014/04/08/half-a-million-widely-trusted-websites-vulnerable-to-heart-bleed-bug.html
4. https://owasp.org/www-project-dependency-check/
5. https://retirejs.github.io/retire.js
6. https://en.wikipedia.org/wiki/Lint_(software)

You can also find the references at **www.mnemonic.no/references-2021**

For more information about mnemonic, visit **www.mnemonic.no**

CONTACT

CORPORATE HEADQUARTERS

mnemonic AS
Henrik Ibsens gate 100
0255 Oslo, Norway
+47 2320 4700
contact@mnemonic.no

STAVANGER

mnemonic AS
Solaveien 88
4316 Sandnes, Norway
+47 2320 4700
contact@mnemonic.no

STOCKHOLM

mnemonic AB
Borgarfjordsgatan 6c
SE-164 55 Kista, Sweden
+46 08 444 8990
contact@mnemonic.se

THE HAGUE

mnemonic
Prinses Beatrixlaan 582
2595 BM Den Haag
The Hague, The Netherlands
contact@mnemonic.io

LONDON

mnemonic Cybersecurity
Level 39
One Canada Square,
Canary Wharf, London E14 5AB
United Kingdom
(+44) 203 973 0036
contact@mnemonic.co.uk

PALO ALTO

mnemonic
470 Ramona Street
Palo Alto, CA, 94301 USA
contact@mnemonic.io

CREDITS

Lead Editor
Rikke Klüver Stenerud, mnemonic AS

Design Lead
Alexandra Stenersen Berg, mnemonic AS

Publication Design
Inventas AS, Oslo

Photo Credits
unsplash.com
stock.adobe.com

The views and opinions expressed in this report are those of the authors and do not necessarily reflect the views of their respective employers.

© 2021 mnemonic AS. All rights reserved. mnemonic and Argus are registered trademarks of mnemonic AS. All other brands, products, or service names are or may be trademarks or service marks of their respective owners.

For more information about mnemonic,
visit **www.mnemonic.no**