A Case Study of the Capital One Data Breach in cyber security

Abstract

In an increasingly regulated world, with companies prioritizing a big part of their budget for expenses with cyber security protections, why have all of these protection initiatives and compliance standards not been enough to prevent the leak of billions of data points in recent years? New data protection and privacy laws and recent cyber security regulations, such as the General Data Protection Regulation (GDPR) that went into effect in Europe in 2018, demonstrate a strong trend and growing concern on how to protect businesses and customers from the significant increase in cyber-attacks. Are current legislations, regulations and compliance standards sufficient to prevent further major data leaks in the future? Does the flaw lie in the existing compliance requirements or in how companies manage their protections and enforce compliance controls? The purpose of this research was to answer these questions by means of a technical assessment of the Capital One data breach incident which occurred at one of the largest financial institutions in the U.S. This incident was selected as a case study to understand the technical modus operandi of the attack, map out exploited vulnerabilities, and identify the related compliance requirements, that existed. The National Institute of Standards and Technology (NIST) Cybersecurity Framework, version 1.1, as a basis for analysis because it is required by the regulatory bodies of the case study and it is an agnostic framework widely used in the global industry to provide cyber threat mitigation guidelines. The results of this research and the case study underline the importance of applying regulatory agencies, establish a more mature cyber security protection and governance ecosystem for the protection of organizations and individuals.

1. Introduction

Technology is nowadays one of the main enablers of digital transformation worldwide. The use of information technologies increases each year and directly impact changes in consumer behavior, development of new business models, and creation of new relationships supported by all the information underlying these interactions.

Technology trends such as Internet of Things, Artificial Intelligence, Machine Learning, Autonomous Cars and Devices, as well as the increasing capillarity of the ever-increasing connection speed, such as 5G (Newman, 2019), result in massive production of information on behavior and privacy-related data from

everyone who is connected. More than 90% of all online data were created within the past two years (Einstein, 2019) and it is expected that these volumes will increase from 33 Zettabytes (ZB) in 2018 to 175 ZB in 2025 (Reinsel, Gantz, & Rydning, 2018).

As the relationships between consumers, organizations, governments, and other entities become ever more connected, there is a tendency for consumers to become more aware of the importance and value of personal information, as well as more concerned about how these data are used by public or private entities (Panetta, 2018). In order to succeed, companies need to earn and keep their client's trust, as well as follow internal values to ensure that clients consider them trustworthy. Based on numerous cyberattacks reported by the media (Kammel, Pogkas, & Benhamou, 2019), organizations are facing an increasing urgency to understand the threats that can expose their data as well as the need to understand and to comply with the emerging regulations and laws involving data protection within their business.

As privacy has emerged as a priority concern, governments are constantly planning and approving new regulations that companies need to comply to protect consumer information and privacy (Gesser, et al., 2019), while the regulatory authorities throughout the world are seeking to improve transparency and responsibility involving data breach. Regulatory agencies are imposing stricter rules, e.g. they are demanding disclosure of data breaches, imposing bigger penalties for violating privacy laws, as well as using regulations to promote public policies to protect information and consumers.

Despite all efforts made by regulatory agencies and organizations to establish investments and proper protection of their operations and information (Dimon), cases of data leak in large institutions are becoming more frequent and involving higher volumes of data each time. According to our research, the number of data records breached increased from 4.3 billion in 2018 to over 11.5 billion in 2019.

There are a number of frameworks, standards and best practices in the industry to support organizations to meet their regulatory obligations and to establish robust security programs. For this research, the Cybersecurity Framework, version 1.1, published by the U.S. National Institute of Standards and Technology (NIST), a critical infrastructure resilience framework widely used by U.S. financial institutions, will be considered as a basis for compliance evaluation.[1] For the purpose of this paper, we selected U.S. bank Capital One as the object of study due to the severity of the security incident they faced in July 2019.

The main research goals and questions of this study are:

1. Analyze the Capital One data breach incident.
2. Based on Capital One data breach incident - Why were compliance controls and cybersecurity legislations insufficient to prevent the data breach?

The result of this study will be valuable to support executives, governments, regulators, companies and specialists in the technical understanding of what principles, techniques, and procedures are needed for the evolution of the normative standards and company's management in order to reduce the number of data breach cases and security incidents.

2. Related Articles

The academic literature related to the objective of this research is limited and, in some cases, outdated, with articles dating from 10 years ago and no connection with the current regulations. The cyberattack trends and the legislation related to data security and privacy have been changing frequently in the past few years. For example, the data leak cases compromising a huge amount of data (millions of data points) have become more frequent recently – in the past 5 years – with a recent trend towards healthcare data leakage and the exposure of huge databases stored in Cloud Computing infrastructures, without the proper access control

[1] NIST published a Cybersecurity Framework in 2014 that provides guidelines to protect the critical infrastructure from cyberattacks, organized in five domains. This Cybersecurity Framework is adopted by financial institutions in the U.S. to guide the information security strategy and it is formally recommended by the governance agencies, such as the Federal Financial Institutions Examination Council (FFIEC).

mechanisms. The frequent updates to the international rules and regulations also contribute to diminish the relevance of older studies.

It is often difficult to get crucial details of the modus operandi of an attack and a list of the compliance controls that failed due to the need to not expose confidential information that could further harm the organization and increase the risk of affecting privacy policies, investigations or confidentiality laws. Furthermore, some regulatory standards do not allow disclosure of details.

Salane (Salane, 2009) indeed describes the great difficulty associated with studies regarding data leaks:

"Unfortunately, the secrecy that typically surrounds a data breach makes answers hard to find. (...) In fact, the details surrounding a breach may not be available for years since large scale breaches usually result in various legal actions. The parties involved typically have no interest in disclosing any more information than the law requires." In fact, it took a detailed analysis of the legal records associated with the data leaks of CardSystems Solutions in 2005 and TJX in 2007, for Salane to identify that both companies were negligent in following the security best practices and the industry's regulatory recommendations. Such records are a rich resource for research, since it provides detailed investigation on the cause of the incidents. However, few incidents have enough technical records available. Hall and Wright (Hall & Wright, Volume 6, 2018) performed statistical analysis of the leaks between 2014 and 2018 and concluded that cyberattacks can happen within any industry: *"It is evident from the research that no industry is immune from the consequences of a data breach."* Hall and Wright also identified that leaks vary over time relative to the type of breach and the type of business affected.

3. Methodological Considerations

This research required the production of preliminary studies that were relevant to this project, allowing the construction of a database with the latest information on data leak incidents that took place between January 2018 and December 2019. This included the identification of relevant information on the type of incidents, who was the target (organization and geography), existence of a technical assessment of the modus operandi of the attacks and the regulations related to the organizations that suffered the attacks.

This research required the availability of technical and trustworthy information regarding the details of the attacks, as well as which regulations were applied at the companies that suffered the data breach. The correlation between the type of data, organizations, country, region, technical details of the attacks, as well as regulations and laws involved are important to answer the key question of the study: Why were or are conformity controls and Cybersecurity laws insufficient to prevent data breaches? Many companies do not disclose the details of the incidents while some will only report and notify clients that their data was compromised, either to comply with regulations, e.g. EU General Data Protection Regulation (GDPR), or involuntarily due to disclosure of details of the incidents by hackers , researchers, the media, or other ways.

One of the greatest difficulties for understanding the modus operandi of the successful attacks that compromised billions of records in the recent years is obtaining detailed information on the attack's vectors, threats, exploited vulnerabilities, technical details of the technological environments and what were the TTPs (Tactics, Techniques, and Procedures) used to compromise the data.

To properly understand the chain of events that led to the incident related to this case study, the MITRE ATT&CK (Adversarial Tactics, Techniques, and Common Knowledge) framework was adopted to help mapping and assessing the TTPs behind each technical step that played a significant role in the success of the cyberattack analyzed.[2] Different from NIST Framework, MITRE ATT&CK is not a compliance and control framework; instead, it is a framework for describing each one of a list of well-known cyber attack techniques, describing their TTPs and related mitigation and detection recommendations. As a result, it helped to determine the security controls that failed or should have been in place to mitigate the attack.

[2] An extensive ATT&CK description is available online at https://attack.mitre.org.

Our background research comprised:

1. 2. This case study containing a detailed analysis to identify and understand the technical modus
3. 4. operandi of the attack, as well as what conditions allowed a breach and the related regulations;
3.1. Technical assessment of the main regulations related to the case study;
Answer to the question: Why were the regulations insufficient to protect the data and what are the recommendations for an effective protection?
Recommendations for regulatory agencies, organizations, and entities.

Technical Criteria for Selection of the Case Study

The first step of the technical analysis was to assess the public records available, if any, about the data leak attacks that were included in the Database of Data Leaks that was built for this study. The objective was to identify the techniques that were deployed in the cyberattack and, as a result, to map the security controls that might have failed.

However, based on the analysis of each case that was mapped in our Database, the public reports for each incident were frequently vague and had little to no details about how the cyberattack took place and how the company was compromised. The greatest challenge in performing the technical analysis stemmed from the lack of detailed reports from trustworthy sources for the majority of the cases that were analyzed. This study considered as trustworthy sources the targeted companies themselves, third party companies involved in the incident investigation and in the response to the cyberattack, information published in legal testimonies and reports provided to regulating agencies, such as the U.S. Security and Exchange Commission (SEC).

3.2. Criteria for regulations analysis (Compliance)

The regulatory scenario is large and permeates several segments in the industry worldwide. When it comes to Cybersecurity, there are strong regulations in the Health and Finance industries (TCDI), among which the most well-known regulations include the Health Insurance Portability and Accountability Act (HIPAA) for healthcare and the Sarbanes Oxley (SOX) and Payment Card Industry – Data Security Standard (PCI- DSS) for the financial industry, in addition to the numerous legislations applicable to a particular country or region such as the General Data Protection Regulation (GDPR) in the European Union, the Brazilian General Personal Data Protection Act (LGPD) and a number of laws in other countries such as the United States. Due to this diversity, it is more productive to select an agnostic framework that is widely used in the industry and offers a mitigation guideline to cyber threats. Thus, the Cybersecurity Framework, version 1.1, published in 2018 by the National Institute of Standards and Technology (NIST) was selected.

3.3. Criteria for Case Study Selection

To choose the Case Study, a survey for a target (company or entity) that suffered a data leak incident between January 2018 and December 2019 was performed under the following two criteria:

1. Had enough technical details publicly available about the incident, and;

2. Public information was available about the regulations to which they were subject and existing compliance report.

Most of the public stories about data leak incidents in 2018 and 2019 did not cover technical details about the incident or had enough information about compliance information on the targeted organization. Usually, press reports only cover superficial information about the type and the extent of the incident. A rare exception was the data breach of U.S. bank Capital One. The incident, which was the result of an unauthorized access to their cloud-based servers hosted at Amazon Web Service (AWS), took place on March 22 and 23, 2019. However, the company only identified the attack on July 19, resulting in a data breach that affected 106 million customers (100 million in the U.S. and 6 million in Canada) (Capital One, 2019). Capital One's shares closed down 5.9% after announcing the data breach, losing a total of 15% over the next two weeks (Henry, 2019). A class action lawsuit seeking unspecified damages was filed just days after the breach became public (Reeves, 2019).

The Capital One case stood out in this research because there is a lot of public information available on the case, since the indictment is available online, including the FBI investigation report (US District Court at Seattle, 2019). In addition, many cyber security consulting companies published blog posts with technical analysis of the incident, such as CloudSploit (CloudSploit, 2019). American journalist Brian Krebs also covered the story, providing some additional technical details (Krebs, 2019). With such amount of information available, it was possible to identify the technical details that describe how the cyber attack took place.

Based on the abundance of details about the incident, as well as the relevant impact to U.S. consumers, the Capital One incident was chosen for the Case Study. In addition, Capital One meets the research criteria since it is an organization working in a highly regulated industry, and the company abides to existing regulations.

4. Hypothesis Procedure

The initial hypothesis of this study was that the current global regulations, normative standards and laws on cybersecurity do not provide the proper guidance nor protection to help companies avoid new data leak incidents.

An additional hypothesis is that the institutions were deficient in implementing and/or maintaining the controls required by existing regulations.

The recent cases of data leaks from large institutions did not result in a quick evolution of the existing standards and cybersecurity policies to minimize or prevent the occurrence of new leaks. For instance, in the Equifax incident in May 2017, criminals stole credit files from 147 million Americans, as well as British and Canadian citizens and millions of payment card records. Equifax will have to pay up to US$ 700 million US dollars in fines, as part of a settlement with federal authorities (Whittaker, FTC slaps Equifax with a fine of up to $700M for 2017 data breach, 2019). The Capital One data breach in 2019 impacted 106 million customers (Capital One, 2019), an initial impact not too much different from the Equifax breach. The editor of news channel TechCrunch, Zack Whittaker, claimed the Capital One data breach was inevitable because probably nothing was done by the industry after the Equifax incident (Whittaker, Capital One's breach was inevitable, because we did nothing after Equifax, 2019):

"*Companies continue to vacuum up our data — knowingly and otherwise — and don't do enough to protect it. As much as we can have laws to protect consumers from this happening again, these breaches will continue so long as the companies continue to collect our data and not take their data security responsibilities seriously. We had an opportunity to stop these kinds of breaches from happening again, yet in the two years passed we've barely grappled with the basic concepts of internet security.*"

5. Case Study: Capital One

5.1. Capital One adoption of technology

Capital One is the fifth largest consumer bank in the U.S. and eighth largest bank overall (Capital One, 2020), with approximately 50 thousand employees and 28 billion US dollars in revenue in 2018 (Capital One, 2019). Capital One works in a highly regulated industry, and the company abides to existing regulations, as stated by them: *"The Director Independence Standards are intended to comply with the New York Stock Exchange ("NYSE") corporate governance rules, the Sarbanes-Oxley Act of 2002, the Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010, and the implementing rules of the Securities and Exchange Commission (SEC) thereunder (or any other legal or regulatory requirements, as applicable)"* (Capital One, 2019). In addition, Capital One is a member of the Financial Services Sector Coordinating Council (FSSCC), the organization responsible for proposing improvements in the Cybersecurity framework, which was selected for this research, and citing the company itself in the appendix published in the NIST website. We also found job advertisements at Capital One's Career website available online in

December 2019 where Capital One was looking for Managers with experience in the NIST framework, which demonstrates that the company had adopted it (Capital One, 2019) (Capital One, 2019) (Capital One, 2019).

Capital One is an organization that values the use of technology and it is a leading U.S. bank in terms of early adoption of cloud computing technologies. According to its 2018 annual investor report (Capital One, 2019), Capital One considers that *"We're Building a Technology Company that Does Banking"*. Within this mindset, the company points out that *"For years, we have been building a leading technology company (...). Today, 85% of our technology workforce are engineers. Capital One has embraced advanced technology strategies and modern data environments. We have adopted agile management practices, (...). We harness highly flexible APIs and use microservices to deliver and deploy software. We've been building APIs for years, and today we have thousands that serves as the backbone for billions of customer transactions every year."* In addition, the report highlights that *"The vast majority of our operating and customer-facing applications operate in the cloud (...)."*

Capital One was one of the first banks in the world to invest in migrating their on-premise datacenters to a cloud computing environment, which was impacted by the data leak incident in 2019. Indeed, Amazon lists Capital One migration to their cloud computing services as a renowned case study (AWS, 2018). Since 2014, Capital One has been expanding the use of cloud computing environments for key financial services and has set a roadmap to reduce its datacenter footprint. From 8 datacenters in 2014, the last 3 are expected to be decommissioned by 2020 (Magana, 2019), reducing or eliminating the cost of running on-premise datacenters and servers. In addition, Capital One worked closely with AWS to develop a security model to enable operating more securely. According to George Brady, executive vice president at Capital One, *"Before we moved a single workload, we engaged groups from across the company to build a risk framework for the cloud that met the same high bar for security and compliance that we meet in our on-premises environments."* (AWS, 2018).

2.2. Technical Assessment of the Capital One Incident

Despite the strong investments on IT infrastructure, in July 2019 Capital One disclosed that the company had sensitive customer data assessed by an external individual. According to Capital One's public report released on July 29, 2019 (Capital One, 2019), *"On July 19, 2019, we determined that an outside individual gained unauthorized access and obtained certain types of personal information from Capital One credit card customers and individuals (...)."* The company claimed that compromised data corresponded to *"personal information Capital One routinely collects at the time it receives credit card applications, including names, addresses, zip codes/postal codes, phone numbers, e-mail addresses, dates of birth, and self-reported income."* The unauthorized access *"affected approximately 100 million individuals in the United States and approximately 6 million in Canada"*, including information from consumers and small enterprises.

According to the FAQ published by Capital One (Capital One, 2019), the company discovered the incident thanks to their Responsible Disclosure Program on July 17, 2019, instead of being discovered by regular cybersecurity operations. The FBI complaint filed with the Seattle court (US District Court at Seattle, 2019) states that Capital One received an e-mail from an outsider informing that data from Capital One's customers was available on a GitHub page (see screenshot extracted from FBI report).

Capital One reported via a press release (PRNewswire, 2019) that some of the stolen data was encrypted but the company did not provide any detail on how it was possible for the attacker to access the information.

"We encrypt our data as a standard. Due to the particular circumstances of this incident, the unauthorized access also enabled the decrypting of data."

According to the FBI investigations, *"Federal agents have arrested a Seattle woman named Paige A. Thompson for hacking into cloud computing servers rented by Capital One, (...). Investigators say Thompson previously worked at the cloud computing company whose servers were breached (...)."* The press soon realized that, according to her LinkedIn profile, Thompson worked at Amazon (Sandler, 2019), indicating that the incident occurred on servers hosted in the Amazon Web Service (AWS) cloud computing infrastructure.

In addition, according to the U.S. Department of Justice (U.S. Attorney's Office, 2019), Paige Thompson was accused of stealing additional data from more than 30 companies, including a state agency, a telecommunications conglomerate, and a public research university. Thompson created a scanning software tool that allowed her to identify servers hosted in a cloud computing company with misconfigured firewalls, allowing the execution of commands from outside to penetrate and to access the servers. The complaint filed with the Seattle court indicates that FBI investigations identified a script hosted on a GitHub repository that was deployed to access the Capital One data stored in their cloud servers. FBI described a script file with 3 commands which allowed the unauthorized access to a server hosted at AWS: the first command was used *"to obtain security credentials (...) that, in turn, enabled access to Capital One's folders"*, a second one *"to list the names of folders or buckets of data in Capital One's storage space"*, and a third command *"to copy data from these folders or buckets in Capital One's storage space."* In addition, *"A firewall misconfiguration allowed commands to reach and to be executed at Capital One's server, which enabled access to folders or buckets of data in a storage space at the Cloud Computing Company"* – according to FBI. FBI adds that Capital One checked its computer logs to confirm that the commands was in fact executed.

After analyzing the records of the Seattle Court, cloud security company CloudSploit published an analysis of the incident in its corporate blog (CloudSploit, 2019), describing that the access to the vulnerable server was possible thanks to a Server-Side Request Forgery (SSRF) attack3 that was made possible due to a configuration failure in the Web Application Firewall (WAF) solution employed by Capital One: *"An SSRF*

[3] Server-Side Request Forgery, (SSRF) is a software vulnerability class where servers can be tricked into connecting to another server it did not intend to, them making a request that's under the attacker's control (Abma, 2017). SSRF flaws occur when an online application requires outside resources enabling an attacker to send crafted requests from the back-end server of a vulnerable web application (O'Donnell, 2019).

attack tricks a server into executing commands on behalf of a remote user, enabling the user to treat the server as a proxy for his or her requests and get access to non-public endpoints."

In his investigation of the incident, American journalist Brian Krebs also concluded that the attacker ran an SSRF attack that exploited a misconfigured WAF tool. Krebs added (Krebs, 2019): *"Known as "ModSecurity,"4 this WAF is deployed along with the open-source Apache Web server to provide protections against several classes of vulnerabilities that attackers most commonly use to compromise the security of Web-based applications."* The diagram that we created (Figure 2) provides a summary of how the vulnerable server was accessed and how the commands were executed by the attacker, leading to the access to sensitive data stored in AWS S3 buckets5 as described bellow.

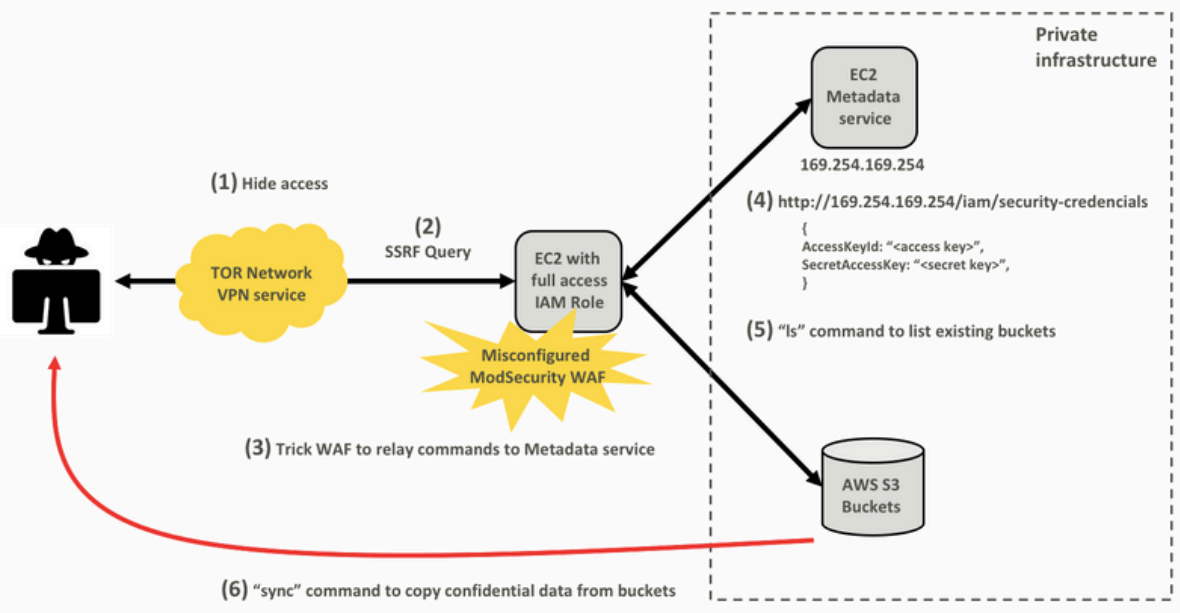

Figure 2: Diagram of the attack: Capital One case study

The reports mentioned above, from FBI, CloudSploit and Mr. Brian Krebs, made it possible to figure out the steps taken during the cyberattack, as presented at Figure 2:

1. The FBI and Capital One identified several accesses through anonymizing services such as TOR Network and VPN service provider IPredator, both used to hide the source IP address of the malicious accesses;
2. The SSRF attack allowed the criminal to trick the server into executing commands as a remote user, which gave the attacker access to a private server;
3. The WAF misconfiguration allowed the intruder to trick the firewall into relaying commands to a default back-end resource on the AWS platform, known as the metadata service (accessed through the URL http://169.254.169.254);
4. By combining the SSRF attack and the WAF misconfiguration with the access to the metadata service containing temporary credentials for such environment, the attacker was able to trick the server into

[4] Modsecurity is a popular open-source, host-based Web Application Firewall (WAF) solution.

[5] Amazon launched its Simple Storage Service (S3) in 2006 as a platform for storing any type of data. Since then, S3 buckets have become one of the most commonly used cloud storage tools.

requesting the access credentials. The attacker then used the URL "http://169.254.169.254/iam/security- credentials", to obtain the AccessKeyId and SecretAccessKey from a role described in the FBI indictment as "*****-WAF-Role" (name was partially redacted). The resulting temporary credentials allowed the criminal to run commands in AWS environment via API, CLI or SDK;

5. By using the credentials, the attacker ran the "ls" command[6] multiple times, which returned a complete list of all AWS S3 Buckets of the compromised Capital One account ("$ aws s3 ls");

6. Lastly, the attacker used the AWS sync command[7] to copy nearly 30 GB of Capital One credit application data from these buckets to the local machine of the attacker ("$ aws s3 sync s3://bucketone."). This command gave the attacker access to more than 700 buckets, according to the FBI report.

The steps described above can be mapped within the specific stages of the MITRE ATT&CK framework, as shown in the table below (Table 1). The ATT&CK framework also describes, for each known attack technique, the main recommendations for mitigation and detection controls that can be used whenever applicable. Therefore, MITRE ATT&CK Framework provides a valuable help by identifying the faulty security controls that made the incident possible.

Stage	Step of the attack	ATT&CK
Command and Control	Use TOR to hide access	T1188 - Multi-hop Proxy (MITRE, 2018)
Initial Access	Use SSRF attack to run commands	T1190 - Exploit Public-Facing Application (MITRE, 2018)
Initial Access	Exploit WAF misconfiguration to relay the commands to the AWS metadata service	Classification unavailable[8]
Initial Access	Obtain access credentials (AccessKeyId and SecretAccessKey)	T1078 - Valid Accounts (MITRE, 2017)
Execution	Run commands in the AWS command line interface (CLI)	T1059 - Command-Line Interface (MITRE, 2017)
Discovery	Run commands to list the AWS S3 Buckets	T1007 - System Service Discovery (MITRE, 2017)
Exfiltration	Use the sync command to copy the AWS bucket data to a local machine	T1048 - Exfiltration Over Alternative Protocol (MITRE, 2017)

Table 1: List of attack steps mapped to MITRE ATT&CK Framework

5.3. Technical Assessment of the Regulations Applied to Capital One

To support this article and the selection of the NIST Cybersecurity Framework both regulatory aspects required by US governance instruments and the best practices were studied.

[6] "ls" is a command available at AWS's command-line interface that list objects and common prefixes under a prefix or all Simple Storage Service (S3) buckets.

[7] "sync" is a command available at AWS's command-line interface that recursively copies new and updated files from the source directory to a specific destination.

[8] MITRE ATT&CK has no specific category that represents the exploitation of a misconfigured cyber security control or tool.

Based on the analysis regarding the regulatory framework applied to Capital One, it was possible to understand the security guidelines provided by Federal Financial Institutions Examination Council (FFIEC), which is a mandatory cybersecurity-related banking regulation in the United States (Miller, 2015).

The FFIEC assumes that the COSO structure (ISACA Control Objectives for Enterprise IT Governance) is the framework elected to support the information security strategy of the financial institutions, associated with the NIST Cybersecurity Framework. According to information made available by Capital One in their investors' webpage (Capital One, 2019), in the scope of Corporate Governance Capital One states that *"The Board of Directors has adopted Corporate Governance Guidelines to formalize the Board's governance practices and to provide its view of effective governance. Our Corporate Governance Guidelines embody many of our long-standing practices, policies and procedures, which collectively form a corporate governance framework that promotes the long-term interests of our stockholders, ensures responsible decision-making and accountability, and fosters a culture that allows our Board and management to pursue Capital One's strategic objectives. The Board reviews and periodically updates these principles and practices as legal, regulatory, and best practice developments evolve."* To map the best-practices that Capital One's professionals follow, we investigated the job descriptions for Capital One's open positions (Capital One, n.d.) to confirm that the abilities and knowledge related to the NIST Cybersecurity Framework are required for those positions. Capital One has strong governance practices regarding cyber security and follows the applied normative frameworks.

While there are numerous regulatory requirements and global standards and best practices covering cybersecurity, this research focused on NIST framework since it is the most comprehensive one.

5.4. Assessment of Technical Controls Versus Normative Standards Applied to the Capital One Incident

This assessment focused on technical controls that could prevent the Capital One data leak incident, according to the incident details published in the U.S. Department of Justice report (US District Court at Seattle, 2019), as described in session ~~5.2~~. In addition, the MITRE ATT&CK framework were used to help map the CSF NIST domains and controls related to the Capital One incident.

For each step performed by the attacker, Table ~~2~~ lists the related technical controls and NIST controls, compromising a total of 61 potential NIST security controls that could have been in place for each stage of the cyber-attack to Capital One. The table supports the conclusion that, possibly, if some of the controls listed here were implemented and operated consistently, the incident wouldn't have been materialized. We will discuss in detail, as examples, two of the control failures in sections ~~6.1~~ and ~~6.2~~.

Deleted:

Deleted:

Deleted:

Deleted:

Stage	Step of the attack	Technical Controls	CSF NIST Failed Controls
Command And Control	Use TOR Network to hide the origin of the attack	Block at Firewall and hosts access from IP addresses from TOR network exit nodes and from malicious proxy server. Alert on IDS/IPS successful access from malicious IP addresses.	**ID.AM-4:** External information systems are catalogued **PR.DS-5:** Protections against data leaks are implemented **DE.AE-1:** A baseline of network operations and expected data flows for users and systems is established and managed **DE.CM-1:** The network is monitored to detect potential cybersecurity events **DE.CM-6:** External service provider activity is monitored to detect potential cybersecurity events **DE.CM-7:** Monitoring for unauthorized personnel, connections, devices, and software is performed **DE.DP-2:** Detection activities comply with all applicable requirements

Stage	Step of the attack	Technical Controls	CSF NIST Failed Controls
Initial Access	Use SSRF attack to well run commands on vulnerable server	Such attack could be mitigated by a well configured WAF and preventive controls, such as periodic vulnerability scanners.	**PR.IP-12:** A vulnerability management plan is developed and implemented **PR.PT-1:** Audit/log records are determined, documented, implemented, and reviewed in accordance with policy **PR.PT-3:** The principle of least functionality is incorporated by configuring systems to provide only essential capabilities **DE.AE-3:** Event data are collected and correlated from multiple sources and sensors **DE.CM-1:** The network is monitored to detect potential cybersecurity events **DE.CM-6:** External service provider activity is monitored to detect potential cybersecurity events **DE.CM-7**: Monitoring for unauthorized personnel, connections, devices, and software is performed **DE.CM-8:** Vulnerability scans are performed **DE.DP-2:** Detection activities comply with all applicable requirements
Initial Access	Explore WAF misconfiguration to send commands to AWS Metadata Service	WAF configuration error could be identified by using preventive vulnerability scan.	**PR.IP-12:** A vulnerability management plan is developed and implemented **PR.PT-1:** Audit/log records are determined, documented, implemented, and reviewed in accordance with policy **PR.PT-3:** The principle of least functionality is incorporated by configuring systems to provide only essential capabilities **DE.AE-3:** Event data are collected and correlated from multiple sources and sensors **DE.CM-7:** Monitoring for unauthorized personnel, connections, devices, and software is performed **DE.CM-8:** Vulnerability scans are performed **DE.DP-2:** Detection activities comply with all applicable requirements

Stage	Step of the attack	Technical Controls	CSF NIST Failed Controls
Initial Access	Get the access credentials (AccessKeyId and SecretAccessKey)	Monitor and audit the use of administrative accounts.	**PR.AC-1:** Identities and credentials are issued, managed, verified, revoked, and audited for authorized devices, users and processes **PR.AC-4:** Access permissions and authorizations are managed, incorporating the principles of least privilege and separation of duties **PR.AC-6:** Identities are proofed and bound to credentials and asserted in interactions **PR.AC-7:** Users, devices, and other assets are authenticated commensurate with the risk of the transaction **PR.IP-1:** A baseline configuration of information technology/industrial control systems is created and maintained incorporating security principles **PR.PT-1:** Audit/log records are determined, documented, implemented, and reviewed in accordance with policy **PR.PT-3:** The principle of least functionality is incorporated by configuring systems to provide only essential capabilities **DE.AE-3:** Event data are collected and correlated from multiple sources and sensors **DE.CM-6:** External service provider activity is monitored to detect potential cybersecurity events **DE.CM-7:** Monitoring for unauthorized personnel, connections, devices, and software is performed **DE.DP-2:** Detection activities comply with all applicable requirements

Stage	Step of the attack	Technical Controls	CSF NIST Failed Controls
Execution	Run commands in the AWS' command line interface (CLI)	Tracking commands on the AWS account	**PR.AC-4:** Access permissions and authorizations are managed, incorporating the principles of least privilege and separation of duties **PR.IP-1:** A baseline configuration of information technology/industrial control systems is created and maintained incorporating security principles (e.g. concept of least functionality) **PR.PT-1:** Audit/log records are determined, documented, implemented, and reviewed in accordance with policy **PR.PT-3:** The principle of least functionality is incorporated by configuring systems to provide only essential capabilities **DE.AE-3:** Event data are collected and correlated from multiple sources and sensors **DE.CM-7:** Monitoring for unauthorized personnel, connections, devices, and software is performed **DE.DP-2:** Detection activities comply with all applicable requirements

Stage	Step of the attack	Technical Controls	CSF NIST Failed Controls
Discovery	Run commands to list buckets in AWS S3	Tracking commands on the implemented AWS account	**PR.AC-4:** Access permissions and authorizations are managed, incorporating the principles of least privilege and separation of duties **PR.DS-5:** Protections against data leaks are implemented **PR.IP-1:** A baseline configuration of information technology/industrial control systems is created and maintained incorporating security principles **PR.PT-1:** Audit/log records are determined, documented, implemented, and reviewed in accordance with policy **PR.PT-3:** The principle of least functionality is incorporated by configuring systems to provide only essential capabilities **DE.AE-3:** Event data are collected and correlated from multiple sources and sensors **DE.CM-7:** Monitoring for unauthorized personnel, connections, devices, and software is performed **DE.DP-2:** Detection activities comply with all applicable requirements

Stage	Step of the attack	Technical Controls	CSF NIST Failed Controls
Exfiltration	Use the sync command to copy data from AWS buckets to local computer	Outbound traffic monitoring	**ID.AM-3:** Organizational communication and data flows are mapped **PR.AC-3:** Remote access is managed **PR.DS-1:** Data-at-rest is protected **PR.DS-5:** Protections against data leaks are implemented **PR.PT-1:** Audit/log records are determined, documented, implemented, and reviewed in accordance with policy **PR.PT-3:** The principle of least functionality is incorporated by configuring systems to provide only essential capabilities **DE.AE-1** : A baseline of network operations and expected data flows for users and systems is established and managed **DE.AE-3:** Event data are collected a sources and sensors **DE.CM-1:** The network is monitored to detect potential cybersecurity events **DE.CM-6:** External service provider activity is monitored to detect potential cybersecurity events **DE.CM-7:** Monitoring for unauthorized personnel, connections, devices, and software is performed **DE.DP-2:** Detection activities comply with all applicable requirements

Table 2: CSF NIST Failed Controls

6. Details of Two of the Failed Controls

Due to the extent of Capital One's data leak incident and the numerous cyber security controls that supposedly failed, two relevant security controls were selected to provide a closer analysis of the security controls whose applicability failed during two relevant steps in the cyber attack chain of events: the privilege escalation that led to the intruder's access to Capital One's server and the data exfiltration.

The existence of technical controls to monitor and audit the use of administrative accounts and to monitor outbound traffic could have prevented the privilege escalation and the data exfiltration, respectively. A deep analysis of remaining mitigation controls is not in the scope of the current paper.

6.1. Case Study: "Obtain access credentials (AccessKeyId and SecretAccessKey)"

Ms. Page Thompson managed to trick the metadata service to request access credentials AccessKeyId and SecretAccessKey (similar to "root access"), which allowed her to run commands in the servers hosted at AWS environment, as explained in section 5.2.

As listed in Table 2, it is expected that the following NIST controls would be able to prevent the attacker to have access to temporary credentials by monitoring and auditing the use of administrative accounts:

- PR.AC-1: Identities and credentials are issued, managed, verified, revoked, and audited for authorized devices, users and processes;
- PR.AC-4: Access permissions and authorizations are managed, incorporating the principles of least privilege and separation of duties;
- PR.AC-6: Identities are proofed and bound to credentials and asserted in interactions;
- PR.AC-7: Users, devices, and other assets are authenticated (e.g., single-factor, multi-factor) commensurate with the risk of the transaction (e.g., individuals' security and privacy risks and other organizational risks);
- PR.IP-1: A baseline configuration of information technology/industrial control systems is created and maintained incorporating security principles (e.g. concept of least functionality);
- PR.PT-1: Audit/log records are determined, documented, implemented, and reviewed in accordance with policy;

Deleted:

Deleted:

- PR.PT-3: The principle of least functionality is incorporated by configuring systems to provide only
- essential capabilities;
- DE.AE-3: Event data are collected and correlated from multiple sources and sensors;
- DE.CM-6: External service provider activity is monitored to detect potential cybersecurity events;
- DE.CM-7: Monitoring for unauthorized personnel, connections, devices, and software is performed;
  DE.DP-2: Detection activities comply with all applicable requirements.

To prevent an attacker from getting the access credentials (AccessKeyId e SecretAccessKey) to perform an exploitation, a set of technical controls are required to restrict the use of user accounts with administrative privileges, as PR.AC-1, PR.AC-4, PR.AC-6, PR.AC-7, PR.IP-1, and PR.PT-3. Monitoring and alerting controls PR.PT-1, DE.AE-3, DE.CM-6, DE.CM-7 and DE.DP-2 would help alarming of any unauthorized access to administrative credential. Therefore, it is very likely that Capital One had insufficient Identity and Access Management (IAM) controls for the environment that was hacked. The periodic review of user and group configuration, in particular the Security Groups, can help ensure that services are not inadvertently exposed, and that the necessary access controls are applied correctly using the principle of least privilege.

6.2. Case Study: "Data Exfiltration"

The Capital One attacker, Page Thompson, ran a sync command on Capital One's server hosted at AWS cloud infrastructure to exfiltrate a large volume of sensitive information by copying data from AWS buckets to a local computer.

As listed in Table 2, the following NIST controls are expected to help prevent data exfiltration by restricting remote access and by monitoring outbound traffic:

- ID.AM-3: Organizational communication and data flows are mapped;
- PR.AC-3: Remote access is managed;
- PR.DS-1: Data-at-rest is protected;
- PR.DS-5: Protections against data leaks are implemented;
- PR.PT-1: Audit/log records are determined, documented, implemented, and reviewed in accordance with policy;
- PR.PT-3: The principle of least functionality is incorporated by configuring systems to provide only
- essential capabilities;
- DE.AE-1: A baseline of network operations and expected data flows for users and systems is established and managed;
- DE.AE-3: Event data are collected and correlated from multiple sources and sensors;
- DE.CM-1: The network is monitored to detect potential cybersecurity events;
- DE.CM-6: External service provider activity is monitored to detect potential cybersecurity events;
  DE.CM-7: Monitoring for unauthorized personnel, connections, devices, and software is performed;
  DE.DP-2: Detection activities comply with all applicable requirements.

The data exfiltration could be prevented by the presence of technical controls to block unauthorized outbound traffic and to monitor outbound traffic from the AWS environment, e.g., by using well known security tools such as Intrusion Detection System (IDS), Intrusion Prevention System (IPS), and Data Leak Prevention (DLP). Such tools could generate alerts that would be subject to specific monitoring.

In addition, Capital One had insufficient controls to alert about commands that users ran in the AWS servers, which avoided them to detect when Paige Thompson ran the specific command line instructions to list and to copy the existing S3 buckets. FBI report shows that Capital One have all the logs regarding the malicious accesses, however the company was unable to detect and to block the access the moment the logs were generated. For instance, AWS has the CloudTrail auditing service, which provides log and monitoring

of the commands ran within the AWS infrastructure.9 A proper monitor and alerting capabilities associated with the commands history would allow the detection of suspicious actions, as the copy of a high number of data repositories.

1. Discussion and Recommendations

1.1. The Compliance impact on cyber security readiness

Throughout the development of this article, the modus operandi of the Capital One attack was understood, as well as the scope of the disciplines contemplated in a mature security framework adopted by the bank. By analyzing the context of the compliance and regulation requirements, one must consider that organizations have the freedom to apply best practice and regulatory controls according to their own interpretation, as well as both technical requirements, business decisions and their risk appetite. In addition, there is a concern on the part of regulatory bodies to allow companies to have the necessary flexibility to adjust the guidelines and controls to fit their particularities, once it follows the proper risk management practices.

In the Capital One incident, for instance, the controls that possibly failed as described in items 5.1 and 5.2 may have been established from a Governance model, but with inadequate parameters compared to the CSF NIST framework. The potential risk for the materialization of cyber incidents lies precisely in this window of opportunity, where an organization is free to interpret the applicability of a compliance control, but the operationalization of such control may not be enough to prevent an incident. In the Capital One case, access management control according to NIST requirement PR.AC-1 was applied, but without considering the premise of least privilege (PR.AC-4), which allowed the attacker to gain the necessary access to exfiltrate the data.

Organizations in general have the challenge to properly establish consistent compliance management across the different teams involved in handling compliance controls, usually organized as “defense lines” across large companies such as Capital One.10 The management and assurance activities performed by the Risk, Compliance, Internal Controls, and both Internal and External Audit teams, have by definition their role in a different time and space where cyber incidents can materialize (see Figure 2).

Figure 2: Cyber Incident window of opportunity

Whereas the Capital One's Technology team (first line of defense) failed to establish proper access controls with less privilege, the time window between identifying and correcting this technical control, either by the second or third line of defense, represents a timeframe where the attacker might exploit. In such scenario,

---

9 CloudTrail provides a history of events for the AWS account activity, including commands ran via the AWS Management Console, SDKs, and the command line tool. (AWS, n.d.)

10 The Institute of Internal Auditors (IAA) adopts the “Three Lines of Defense Model” to explain the relationship between the teams involved in the ownership and responsibility for operating risk management and control (Chartered Institute of Internal Auditors, 2019).

any of the teams (lines of defense) would be able to identify and to demand the correction of the weak control prior to exploration. Based on the public information available regarding this incident, it was unable to corroborate the position of the Capital One auditors. Continuous audits through online compliance monitoring can assist with timely decision making and mitigate the risk of this kind of incident occurrences.

1.2. The cyber security GAP between Governance, Management and IT

Capital One's digital transformation journey to migrate its entire technology platform to the cloud presented a well-planned strategy. They hired talent engineers, invested financially on multiple fronts, hired a renowned CISO, and even supported AWS developing a series of tools like Cloud Custodian to have a portal where they can see and measure compliance in the entire fleet of services throughout their complex, multi-account environment.

But, all of these actions were not sufficient to prevent the data breach incident. For a growing number of companies around the world, it's now commonplace to store sensitive data on AWS's public cloud Simple Storage Service (S3). While reliable, S3 also pops up in the news very often because of the exposure of unsecure databases, when misconfigured security settings leave data publicly exposed to the internet. AWS has added services designed to prevent security misconfigurations, in addition to strong defaults and pop-up warnings which note if a bucket has public settings. Regarding the Capital One incident, AWS said its cloud unit that stored the data was not compromised in any way. Instead, it attributed the breach to a "misconfiguration" outside of the cloud. Capital One attributed the problem to an error in its own infrastructure (Henry, Capital One customer data breach rattles investors, 2019). The misconfiguration issues for some reason have not been detected and avoided by the security controls that Capital One claims to implement. A human error might be a cause. Even before the incident, some Capital One cyber staff raised concerns about employees morale: *"employees raised concerns within the company about what they saw as high turnover in its cybersecurity unit and a failure to promptly install some software to help spot and defend against hacks (...). While the bank was generous with cybersecurity funding, the unit struggled to stay within its budget last year.... This year (2019), budget issues have continued and possible money-saving measures, including staff cuts, have been discussed... Routine cybersecurity measures to help protect the company sometimes fell by the wayside.... the bank around late 2017 bought software from a company called Endgame to improve its ability to detect hacks... More than a year after buying the software, Capital One still hadn't finished installing it... The issue was flagged to Mr. Johnson (CISO), the bank's internal auditors and others, according to one of the people. It couldn't be determined how they responded... The bank's "red team," an internal group intended to find vulnerabilities in the firm's security, once broke into the private elevator to executive floors, a move some employees involved thought would have crossed the line at other firms... Sometimes the broader tech-centric culture of the firm could complicate security... Technology employees had at times been given free rein to write in many coding languages — so many that it made it harder for the cybersecurity unit to spot problems."* (Andriotis & Ensign, 2019) While cybersecurity skills are in high demand and companies are ready to hire top talent, weak leadership and a toxic culture can quickly lead to employee retention issues. This is not a technological risk, but a management risk that can impact critical actions of an organization. In addition to the many negative consequences for the image and stock after the incident, Capital One also changed its chief information security officer out of the role. No other consequences could be identified for other compliance, audit or technology employees. Even with these "misconfiguration" and management issues, Richard Fairbank, CEO at Capital One said:

"We remain absolutely committed to our digital strategy and our technology transformation, and the public cloud is an essential element of that strategy."

1.3. Recommendations to mitigate and strengthen the standards based on Capital One case study

1.3.1. To avoid the improper adoption of compliance controls

Highly regulated industries, such as health and financial institutions, make an effort to refine their compliance guidelines and to limit the companies' autonomy in pursuit of an increasingly uniform and collaborative environment. As an example, the Federal Financial Institutions Examination Council (FFIEC) in the United States published a detailed set of controls, the so-called Technology and Security Booklets. This posture can be beneficial for companies as the operation will have more inputs to enforce controls, as auditors will have more assertive audit criteria to achieve the objective.

1.3.2. To keep the controls relevant as the technology evolves

Compliance standards, legislations and regulations demand a long-term effort from the industry and regulatory bodies to be developed and updated in a regular basis. As a consequence, keeping up with the constant technological changes is a major challenge for the applicability of compliance controls. In the Capital One case, and most of other data leak cases in 2018 and 2019, existing security controls applied to Cloud Computing storage properties were not properly configured to prevent the access and exfiltration of sensitive information. As shown in Table 2, the controls proposed by CSF NIST could prevent the security incident from materializing if properly applied.

1.3.3. Multidisciplinary Skills

The technical qualification of IT and Compliance professionals is an important point to consider. By working with modern and advanced technologies, in an interconnected online business, employees require multidisciplinary skills and frequent training. In addition, even professionals with extremely technical roles as web developers and IT architects, need to improve their security and governance skills in order to properly apply such controls to their context, in the same sense that governance professionals must be able to understand the technological requirements applied to their IT environments. In addition, companies have to establish a governance structure that establishes the approval and action mandates, so that decisions are made timely.

1.3.4. How to protect a Storage (S3) Cloud Environment

Our research shows that many recent incidents are related to misconfiguration in cloud storage, for example, AWS S3 buckets. Some security controls to mitigate this type of vulnerability in AWS include:

- **Know the infrastructure and know which users can access what and why;**
- **Apply a Principle of Least Privilege. Use AWS Identity and Access Management (IAM) user policies to specify the users that can access specific buckets and objects;**
- **Separate resources and do not mix private and public data within an S3 bucket;**
- **Manage all entities and enable blocking public access;**
- **Keep the infrastructure up to date;**
- **Amazon offers a WAF solution which integrates with CloudFront and blocks suspicious requests before they reach the servers;**
- **Monitor the S3 buckets using products like AWS Config, AWS Cloudtrail and Lambda. Enable email notifications from trusted Advisor to get notified of unintended changes to the bucket policies and bucket ACLs. Run the Amazon S3 Bucket Permissions check;**
- **Follow all the best practices as NIST CSF and the vendors recommendations.**

1.3.5. The need to manage the compliance window

The time lapse between a compliance control being evaluated, implemented and audited represents an important element to be considered by organizations wishing to enhance their cyber defense capabilities. Organizations can benefit from filling the gap with ongoing monitoring and auditing activities, by increasing the monitoring of their operation, from technical infrastructure (done by Network Operation Centers – NOC) and security-related incidents and vulnerabilities (managed by the Security Operation Center - SOC), with the regulatory and governance aspects, building a Governance Operation Center (GOC). Such approach would help to continuously measure the efficiency of the existing compliance controls in real time, as well as being fertile ground for analytics initiatives given the amount of multidisciplinary data.

2. Final considerations

The study of the Capital One incident showed that the company failed to implement proper security controls. It also demonstrated that the NIST Framework would have been sufficient to mitigate the incident, if there were enough compliance controls in place to identify the unauthorized access and data exfiltration during the entire chain of events.

The many cases of information leak incidents show that companies worldwide are not properly adapted to use and to manage the security of new cloud computing environments, even when compliance controls do exist and vendor guidance is in place to provide support to companies and secure their environments.

From a global perspective, regulatory agencies must ensure that proper compliance frameworks and regulations are in place to support local companies. For example, in Latin America the absence of legislation enforcing the use of well-stablished standards such as the NIST or ISO frameworks means that companies based in these regions are not required to implement such controls that would prevent further incidents - except when the organization itself takes the initiative to apply such frameworks on their own. In Brazil, for instance, local banks have to comply with cyber security controls enforced through Central Bank Rule 4658, and also by existing laws such as LGPD. However, these standards lack controls as complete and comprehensive as NIST. That being said and considering that many companies operate globally with customers and suppliers (supply chain) potentially anywhere in the world, such as the new global banks, we recommend that companies adopt global governance frameworks that feature cyber security controls capable of addressing new technologies. In an increasingly connected world that breaks down continental barriers, a weaker security standard tends to be compromised and will, therefore, contribute to the compromise of other organizations even if such organizations follow stronger standards. In other words, a local failure can impact everyone in the industry, which is the reason why we need a global policy for data protection.

3. Future work

Through this research, we highlighted the need to better understand and discuss the information about data leak incidents. We recommend extending this study by analyzing other incidents.

Further work might include the analysis of other legal standards that Capital One failed to comply with in order to prevent the July 2019 data leak incident, understanding how compliance and governance by other regulating agencies are applicable, the role of cyber insurance, and how to manage the culture and management to maintain the best talents. Audit reports, whenever they exist, are confidential and not accessible to the public, making it difficult to understand and analyze the effectiveness of existing controls. A global perspective must be included to help analyze which countries or geographies adopt the NIST framework, ISO standards and/or relevant local regulations. A discussion on how specific countries apply well-known compliance standards to local organizations will help in understanding the extent of cyber security governance in a global and local level.