A Study of the Implications of Logistical and Social Issues in the Security and Privacy of Cloud Computing as they Relate to Public, Private, and Classified Workloads

A Thesis Prospectus Submitted to the

Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements of the Degree
Bachelor of Science, School of Engineering

,

Andrei Stan
Fall, 2020

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Signature _____ Date _____

Andrei Stan

Approved _____ Date _____

Sharon Tsai-hsuan Ku , Department of Engineering and Society

STS Prospectus

Introduction

Cloud computing is a large part of daily life and consists of a large part of even the professional computing and workloads done in the world today. In the private sector as well as the much higher security demanding public sector, the security, privacy, and handling of sensitive data in the cloud and in these sometimes classified workloads is a persistent and difficult issue. This is a somewhat new issue that has difficult implications about how and where to handle the data at play. This novelty has led to rapid handling on both sides of the industry, with public opinion as well as policies and protocols evolving quickly. Legislation has attempted to keep up with this rapid growth and, even as it continues to grow, this issue of privacy and security persists around cloud adoption and development. This study will attempt to concretely analyze and identify these issues as well as identify the most key of these and put forward potential plans for comparing and handling said issues.

Research Question

Cloud computing offers many benefits over traditional workloads. However, security in the cloud, in many cases, is very different to how it is handled traditionally. User data and processing is done in data centers owned by service providers like Amazon Web Services(AWS) and Microsoft which presents logistical issues like physical data center security and access as well as sociotechnical issues like the proper use of customer data and its ownership.

Security needs always exist, but particularly in the public sector, intelligence agencies or government contractors need an assurance of the same level of high security that they are able to achieve with their on-site infrastructure. In order to match these security needs, service providers may have accreditations or certifications that, at a technical level, show qualification to handle potentially confidential or highly classified workloads.

Even when the technical qualifications are met, social disagreements may occur. Individuals on both sides may have negative predispositions which cause them to be particularly skeptical or concerned about the safety and privacy of their data. Assurances like highly restricted physical access to data centers may even serve to deter certain users, especially those with prior distrust towards service providers or the industry as a whole.

Documentation from service providers and accreditors can be used to understand the technical aspects of different levels of workload security. This documentation can then be compared with records of high profile failures in order to potentially identify core issues that lead to some of the problems with security and privacy in cloud computing.

Literature Review

Cloud benefits and tech aspects were found to be the most important when considering cloud adoption(Khayer et al., 2020) but resistance to change and computer self-efficacy are also important. This is expanded by describing a divide between, at least in Europe, areas of differing economic and technological development in terms of factors that influence cloud adoption. Arvanitis, Kyriakou, & Loukis (2017) found the less developed Southern Europe to be concerned with financial issues while the more developed Northern Europe to be more concerned with factors involving business efficiency, like increased innovation and capabilities.

Security was found to be an issue for cloud adoption for students in educational environments as well (Arpaci et al., 2015). It was also found that perceived security is the deciding factor when considering cloud adoption and use. This can be influenced by assurances but initial perception was found to have a bigger effect. This is consistent with the findings of

Shah et al. (2014) in that website design had the biggest effect on perceived security as opposed to what would be expected to have a larger impact in terms of assurances (both external and internal) and the solutions actually available for confidentiality.

This social perception was also found to be mainly affected by education about cloud terminology and concepts, at least in the Jordanian government context. Different countries will behave differently but a lack of education in this context led to some subjects confusing terms as well as having opinions that did not match real world facts (Alkhwaldi, Kamala, & Qahwaji, 2019).

These concerns do seem to have basis however, according to King & Raja (2019), who highlighted the need for regulatory reform as far as the foundations of security and protection of sensitive data in cloud computing. They argue that current definitions of sensitive data are not adequate for proper protection of users or for the success of cloud service providers.

Contrasting these perceptions, however, cloud adoption has several methods for handling security and privacy issues. These methods are well documented in the context of AWS and Azure (Rath et al., 2019) as well as many other providers. The benefits of cloud are also well known but the issues seem to center around security once again, at least in the context of e-government, according to Alshomrani & Qamar (2013).

On the opposite side, Hughes et al. (2019) propose a new idea of not segregating the use of governments. They suggest the lack of potential for things like GovCloud in terms of growth and propose the use of the public, widely available and continuously growing cloud using special security techniques in order to place the responsibility back in the hands of customers with respect to government or classified workloads.

The findings of Norris & Reddick (2012) show that the transition to cloud may not be at a point of revolution and likely never will be. They found the transition of governments to electronic domains to be more of an incremental shift rather than a revolutionary one in the context of local e-governments. This can be interpreted as either agreeing or disagreeing with the findings of Hughes et al. (2019). Both a transition to segregated clouds like GovCloud and the idea of using public cloud in specialized manners but in mass can be considered examples of revolutionary changes and both can also be considered incremental as both are examples of using what is already available and slowly and effectively leveraging current resources for government purposes. This would depend on the context of the use and again, is very specific to each user's perceptions, needs, and capabilities.

STS Framework and Method

The issues regarding sensitive cloud data can be labeled as ones of society as well as technical. The frameworks chosen for analysis of these issues are relevant to this type of classification and attempt to understand these issues. Hughes' perspective was used in order to understand the sociotechnical aspect behind the information. A SCOT analysis was used to target the specific social groups at play as well as their interests and conflicts. Finally, ANT analysis was used to combine the issues and identify the actants at play as well as their translation as part of the identified ANT network.

<u>Hughes' Perspective Analysis:</u>

The builders behind cloud storage are mainly the developers and engineers behind the specific services. They obviously work on the programming, hosting, and software-level security of the systems. Additionally, there are support teams and maintenance personnel that are responsible for running these services and catering to customers. Pertinent specifically to the topic, the maintenance personnel are often a majority of the few people that have access to the

physical data centers that house the cloud storage. In addition to maintenance staff, third party auditors also play a role as "system builders" by providing a security safety net and potentially reassurance of the security of the services to customers.

In the case of AWS, the most important voice is the customer, per their "working backwards" process. Customer demands are said to be directly implemented in the majority of cases. Politically, there are also many customers in the public sector that have a big stake in how these services are developed and maintained. Increased security is required for customers in the intelligence and defense communities. These customers can be seen as major stakeholders, especially in cases like the dispute over the $10B Pentagon JEDI contract between AWS and Microsoft.

There are also differences that cause customer preference between providers. In AWS's case, being the biggest provider, their main selling point is the high availability offered by their infrastructure. In Microsoft/Azure's case, the main selling point is that they are able to offer their services at a much cheaper price due to hybridization integrated into the native Windows environment. Smaller cloud providers also specialize in certain services, with the example of Oracle being one of the most experienced database service providers or IBM having spent the majority of their resources recently specializing in artificial intelligence and machine learning services.

With cloud computing being at the forefront of current computing technology, there are not very many reverse salient aspects that are due to age. Many can be identified as "growing pains" like the slow adoption of the services themselves by customers due to a large gap in expertise between providers and customers. Once again though, the security of the cloud is a huge issue. The fundamental idea of storing private data or records on external hardware, potentially accessible by other people outside of a specified group is something that cloud computing has no clean solution for. While governance, policies, and certifications attempt to ease some of these concerns, there can never be any assurance other than trust and laws or regulations that can completely remove these concerns.

The public sector adaptation of this technology is the biggest driving factor in the adoption of cloud services. Security is a concern but the example of intelligence agencies and defense organizations adopting these services for mission critical and highly secure workloads serves as good motive for other, private sector or individual customers to put their trust in the security behind cloud services.

SCOT Analysis :

In the context of cloud storage security, there are only a few relevant social groups. Users make up the entirety, with the different types having different needs or regulations governing their decisions. In the private sector, users range from single users like students to multi billion dollar, Fortune 500 corporations, like Capital One, General Electric etc. In the public sector, users include contracting firms, universities, and even defense organizations like the FBI and CIA.

Cloud security considerations vary among these different types of users, ranging from little to no consideration for smaller customers in the private sector like students to a top priority, such as for defense agencies. These users that require a high level of security even go as far as operating completely disconnected from the internet, through SIPRnet and JWICS. For customers that operate on the internet, namely users in the private sector, security considerations also widely vary, especially in the case of intellectual property being stored in cloud services by larger companies and even by some individual users.

These requirements can sometimes be negotiated. Especially in the private sector, some users are willing to compromise on security in order to user cloud services. This is not, however, common. Most users that do have security requirements are not at all willing to budge, and cloud service providers often know this. Most services are built around the idea that security is the highest priority. This can especially be seen in the implementation of public sector cloud technologies, with many data centers being explicitly dedicated to certain customers and solutions like dedicated servers/storage being offered to allow for stricter control of data access.

Closure mechanisms do exist but the fact still stands that the services that cloud providers offer run through non-privately owned hardware. A physical person or group of people not employed by the customer has to have access to the physical hardware in order to guarantee that the services continue to run. Third party auditors exist to give the customers a sense of relief in that the physical hardware and data centers are inspected and certified to meet certain security standards and compliances like FedRAMP, ITAR, FERPA etc. There are also those data centers in which, if required by the users' workloads, all personnel are required to meet certain clearances, such as secret or top secret, in order to even be allowed to access the data centers.

ANT Analysis:

Specific to the concept of limited physical data center access as well as societal perception of cloud adoption, a human-non human network is created. This network centers around the idea of careful control of processing and data specific to workloads with highly sensitive data directly conflicting with the policies in place at most cloud providers that are created for the very purpose of higher security and privacy of said sensitive data. While on one side, the idea of limiting physical data center seems like the perpetrator in that it may cause a lack of trust, the societal aspect serves as an issue in this network as well as it may have a non-pragmatic view or perception of cloud and of what actually happens "behind the scenes" in said data centers.

There is some successful translation here because the concept of a black box of computers and servers that handle workloads in highly classified environments directly contrasts the idea of careful handling by humans of the data in these workloads. When an issue arises over this generalisation of data centers and data handling, the societal aspect of the human stakeholders and decision-makers in cloud adoption can now be seen as that of an obstacle in the way of adoption and further development while the actual policies and handling of the data center access policies can be seen as a malicious tactic to either steal customer data or sell it for profit. It is likely only in these situations, when specific circumstances like negative predisposition towards cloud adoption as well as malicious intent and actions on the side of the physical data or processing side that this translation can come across.

The concepts can be somewhat explained by the actant network. They are symbiotic characteristics where the societal view and general uneasiness about privacy comes as a result of previous issues and failures that have caused people to be more aware and place a greater value or importance on their sensitive data. This very same uneasiness has led to the development of such strict measures as the level of control of the physical data centers or the encryption and control of such things as access keys. These measures, while intentions may be good behind their establishment, can directly contrast societal opinions and create a positive feedback loop of distrust and stricter control further driving apart the middle ground solution for the proper sensitive data handling procedures.

Methods for Data Collection:

In order to attempt to figure out a solution to the issues behind successful cloud adoption, a document analysis will be used. In this situation, surveys, interviews, and participant observations would not be viable as the nature of the issues specific to this project are involving classified information that stakeholders would not be allowed to communicate about.

Instead, a document analysis will be used to attempt to understand what the key, concrete issues are. On the technical, data storage side, documents from cloud providers and documentation of the services as well as technical papers concerning the use of said services will be analyzed to try to understand what the exact policies and technicalities are and what loopholes or issues may arise from them.

On the societal side, news articles as well as documentations and technical analyses of breached can be analyzed to understand the motivations and possible misunderstandings behind negative predisposition towards cloud adoption. In addition, case studies available both publicly from cloud storage providers as well as in the form of more technical analyses can be used to understand how failures are handled and who, generally, receives the blame or is the subject of the majority of negative societal reactions.

Once both sides are properly analyzed and all details are understood, a few solutions can be theorized about how to handle these issues and what possible next steps could be in improving this lack of trust regarding sensitive information. Should it be the case that no obvious possible solutions exist, more analysis could be done on both sides to try to find points of compromise or ideas that could be the source of new topics of discussion or a dive into deeper issues than just predisposition and the causes behind it or the types of assurances that are currently in place.

Because the primary contributor has some experience with working with these cloud services, it will be important to address this inherent bias in the study since it may influence their decision-making. This is also true of the contributor's own personal feelings and opinions towards the current cloud computing industry. However, as this study is strictly based on analysis of resources and does not involve any test subjects or experimentation, the impact of the inherent biases will be less significant.

Timeline

The project will be conducted throughout the course of the Spring 2021 semester. An understanding of the issues faced in cloud computing data privacy and security, as well as appearances in media and general public attitude will be analyzed in the earlier parts of the semester. These data will then be reduced down to key issues/attitudes and then recorded as issues to specifically target when looking into the technical procedures and documentation on the technical side. This secondary analysis will be performed in the later part of the semester, leaving time for a proper analysis of both sides as well as time to attempt to come up with a few possible solutions to the key issues identified earlier in the semester.

Conclusion

Public opinion of cloud computing handling of sensitive data is a pressing issue that currently faces the industry and the world. There are many different factors that affect not only public opinion and sentiment but also the process by which providers handle sensitive data as well as legislation regarding said data. Because of this large number of factors, it is not expected that there will be one, concret, obvious solution to the multitude of issues, nor that there is an obvious direction of improvement.

Instead, this project will attempt to identify, concretely, the key issues that face this industry and public sentiment and opinion and come up with ways in which these can be

addressed directly for the future in an attempt to push the handling of these complex issues in the right direction.

Bibliography

Alkhwaldi, A., Kamala, D. and Qahwaji, P., 2019. Security Perceptions In Cloud-Based E-Government Services.

Alshomrani, S. and Qamar, S., 2013. Cloud Based E-Government: Benefits And Challenges. [online] Ijmse.org. Available at: <http://www.ijmse.org/Volume4/Issue6/paper4.pdf>.

Arpaci, I., Kilicer, K. and Bardakci, S., 2015. Effects of security and privacy concerns on educational use of cloud services. Computers in Human Behavior, 45, pp.93-98.

Arvanitis, S., Kyriakou, N. and Loukis, E., 2017. Why do firms adopt cloud computing? A comparative analysis based on South and North Europe firm data. Telematics and Informatics, 34(7), pp.1322-1332.

Corstorphine, A. (2020). The Capital One Data Breach a Year Later: A Look at What Went Wrong and Practical Guidance to Avoid a Breach of Your Own - Security Boulevard. Retrieved 14 February 2021

Hughes, J., Munson, C., Schear, N., Patel, R. and Kalke, M., 2019. Classified As A Service (ClaaS). [online] Apps.dtic.mil. Available at: <https://apps.dtic.mil/sti/pdfs/AD1100938.pdf>.

Khayer, A., Talukder, M., Bao, Y. and Hossain, M., 2020. Cloud computing adoption and its impact on SMEs' performance for cloud supported operations: A dual-stage analytical approach. Technology in Society, 60, p.101225.`

King, N. and Raja, V., 2012. Protecting the privacy and security of sensitive customer data in the cloud. Computer Law & Security Review, 28(3), pp.308-319.

F. Massacci, T. Jaeger and S. Peisert, "SolarWinds and the Challenges of Patching: Can We Ever Stop Dancing With the Devil?" in *IEEE Security & Privacy*, vol. 19, no. 02, pp. 14-19, 2021. doi: 10.1109/MSEC.2021.3050433

Mthunzi, S. N., Benkhelifa, E., Bosakowski, T., Guegan, C. G., & Barhamgi, M. (2020). Cloud computing security taxonomy: From an atomistic to a holistic view. Future Generation Computer Systems, 107, 620–644. https://doi.org/10.1016/j.future.2019.11.013

Norris, D. and Reddick, C., 2012. Local E-Government in the United States: Transformation or Incremental Change?. Public Administration Review, 73(1), pp.165-175.

Rath, A., Spasic, B., Boucart, N. and Thiran, P., 2019. Security Pattern for Cloud SaaS: From System and Data Security to Privacy Case Study in AWS and Azure. Computers, 8(2), p.34.

Shah, M., Peikari, H. and Yasin, N., 2014. The determinants of individuals' perceived e-security: Evidence from Malaysia. International Journal of Information Management, 34(1), pp.48-57.