[1] S. Kesavan

[2] S. Srinivasan

A Trusted Execution Environment-Based Data Aggregation System that Preserves Privacy for Fog / Cloud-Enhanced IoT Applications

Abstract: - The Gadgets for the Internet of Things are growing more and more common in our day-to-day activities. This is because they present a plethora of opportunities developers and the sector to create applications, leveraging their low cost, low size, and high connection. Because many these programs are similar in that they gathered information using cloud servers and fog because of the computational limitations among the peripheral gadgets, concerns regarding The secrecy and privacy of The information these gadgets produce is increasing i tandem with the ongoing rise in their usage. To provide clients with personalized services, data collected from several devices is combined and evaluated using cloud and fog environments. Given that IoT devices have the ability to gather sensitive user data, including behavioural and personal This study outlines current efforts to develop a generic data aggregation system that makes Utilizing Secure Execution Environments (TEEs) to protect user privacy and data by enabling the interpretation of various data and the execution of complex computations, such as the application of machine learning techniques. We describe the design of the system, our preliminary findings, and the next steps to put our plan into practice and validate it.

Keywords: SGX, sensitive data, data processing, user privacy, security, and confidentiality

I. INTRODUCTION

The idea of "connected objects," or "IoT," has attracted a lot of attention recently because it can be used to investigate new industries including industrial control, logistical, medical, and surveillance more by linking numerous objects. Even though Things connected to the Internet (IoT) have many benefits, such as affordability, adaptability, and remote management, their capabilities frequently stop them from executing more complicated tasks because of limitations in battery life, memory, and processing. These limitations can be circumvented by integrating with cloud environments [1].

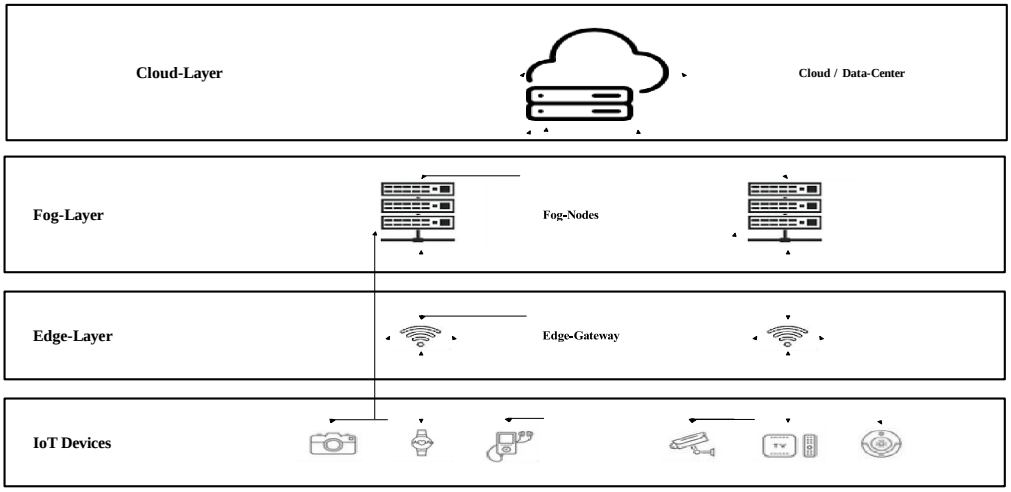
Resource pooling is a fundamental aspect of the cloud computing architecture, wherein workloads related to several users, referred to as tenants, are commonly consolidated into a single pool of physical resources. Elastic scaling capabilities of cloud computing allow for resource expansion or contraction in response to user demand with little administration work or contact between service providers [2].

As alternatives to cloud computing, there is fog and edge computing. The gateway is situated adjacent to the Internet of Things devices In the case of edge computing, local data processing takes place. In [3]. For data processing and storage, local servers may be used. when greater processing power is needed and response times are tight. This condition, which reduces costs and delays by using cloud processing and communication, is known as fog computing [4]. Furthermore, data processed at the fog and edge is sent to the cloud for analysis and system-to-system communication. Figure 1 shows the link between gadgets, cloud, fog, and edge environments.

[1] Research Scholar, Department of Advanced Computing Sciences, AMET University, Chennai, India.
Email: kesevan2007@gmail.com

[2] Professor, Department of Advanced Computing Sciences, AMET University, Chennai, India.
Email: srinikcgmca@gmail.com

Copyright © JES 2024 on-line : journal.esrgroups.org



Typical examples of Internet of Things applications are displayed in Figure 1. The possible communication routes between the devices and the fog, cloud, and edge layers are depicted by the arrows.

There is more processing power available in cloud and fog environments for compiling and analyzing information generated by IoT devices. IoT applications benefit greatly from this integration, although some worry about privacy of data and confidentiality that must be considered account because Numerous IoT applications address sensitive data, like financial, medical, and private data, which could be enticing and useful to hackers.

As such, safeguarding customer information and privacy is one of the most important concerns at the moment. Since data generated by Internet of Things devices can reveal a great deal about their users, it needs to be handled properly. One way to handle In fog and cloud environments, it is advised to use privacy-preserving data aggregation techniques for sensitive data; however, most of these algorithms only offer one type of aggregation for homogeneous IoT devices and do not allow aggregating data from heterogeneous devices [5]. Furthermore, security is an important consideration when combining data [6].

The goal of our effort has been to fill a vacuum in the literature since existing approaches do not address aggregating strategies that protect privacy for a variety of devices and data. Swarms are large, dynamic, self-organizing networks made up of the heterogeneous devices seen in many Internet of Things systems [7, 8]. For creating and evaluating privacy-preserving strategies for heterogeneous data types, there isn't a publicly accessible tool.

In this work, we offer a novel fog data collection strategy with privacy-preserving limitations from many IoT devices and cloud environments using Intel SGX, a Trusted Execution Environment. This technology development is widely used in cloud- and fog-enhanced Internet of Things applications, providing safe techniques for processing and storing sensitive data that guarantee confidentiality and integrity [9].

The remainder of the essay is structured in this manner: The system architecture of the work's recommended fix is shown in Section V; Part VI describes how a proof of concept is put into practice, and piece VII, the final piece, describes the next steps in our research; Part III describes the challenges associated with data aggregation with privacy preservation, and relevant literature is covered in Section IV. The ideas behind Intel SGX technology

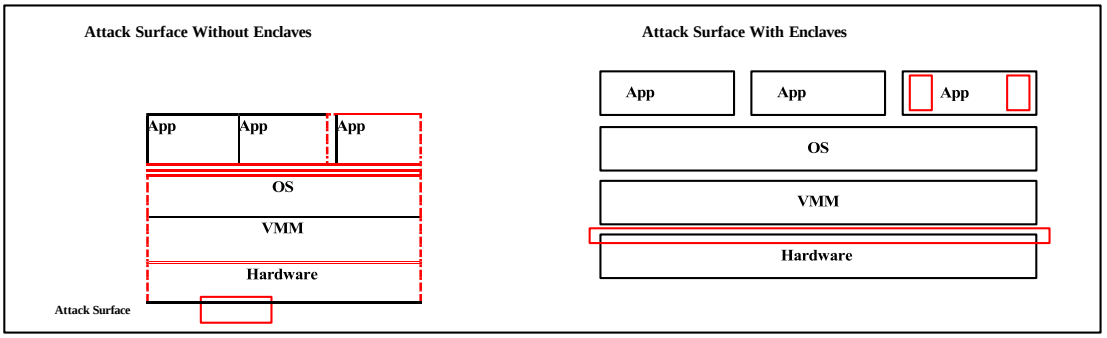
II. Intel SGX

Intel CPUs now have new instructions called Extensions for Software Guard (SGX) that allow programmers to put sensitive code and data in a safe area known as an enclave. Enclaves maintain secrecy and integrity in a memory area that is encrypted. Furthermore, primitives for access control are present in this memory, preventing hostile programs from accessing any data, including sensitive ones like the operating system, BIOS, and hypervisor [10].

The reduction of restricting the Trusted Computing Base (TCB) to just one piece of gear and piece of software

component is the primary objective of the SGX, as illustrated in Fig. 2. Because of this, software is less susceptible to compromise due to a decrease in surface assault.

When an enclave is loaded from memory, SGX can detect if any of its components have been changed without authorization. If so, it will prevent the loading of that enclave.



Furthermore, the enclave is connected to the program that established it, allowing for analysis and study [10].

Enclaves are able to communicate securely and exchange information An enclave can demonstrate that it is genuine, unmodified, and correctly loaded by employing an attestation technique to verify each other. During the Diffie-Hellman agreement method, a symmetric key is defined by local attestation when both enclaves are operating on the same platform. The hardware then verifies the key. There is also remote attestation, which enables the establishment of secure networks for the connection of external devices. For an enterprise to be prohibited from disclosing sensitive information to any entity other than the verified enclave, Establishing and connecting a secure route to the attestation procedure is necessary [12].

Moreover, SGX offers means for securely storing confidential information in permanent memory through the sealing procedure. A distinct key, known as a sealing key, which the CPU and enclave produce identity can be requested by any enclave from the CPU. Only the identical enclave and CPU may unseal the data, according to the data sealing functionality [12].

III. DATA AGGREGATION: PRIVACY-PRESERVING

Aggregation refers to the process of gathering data and arranging it for presentation. To create a summary for data analysis, information can be taken from many different sources. Since it can provide thorough and trustworthy data that improves the dependability and quality of information systems, data aggregation is an impressive area of study.

It is well known that datasets collected from various devices for Internet of Things applications often contain highly connected and duplicated data. As such, data aggregation—a common feature of Internet of Things systems—becomes an effective means of combining such redundant data and delivering information.

Furthermore, data protection is essential because to the vast quantity of devices that gather and handle potentially sensitive data, particularly for Internet of Things applications that are enabled by fog and cloud settings and interchange obtained data with other businesses., forwarded to other suppliers, and may even be outsourced [13, 14].

IV. LINKED WORK

A number of strategies have been used in the literature by authors who have written about privacy-preserving data aggregation. Because many of the recommended ways are unable to manage heterogeneous data or hybrid devices, Lu et al. [5] call attention to the fact that many Internet of Things applications cannot use them. By combining data aggregation through Paillier homomorphic encryption with one-way hash chain techniques for authentication, the authors provide a remedy for this problem. In addition, [15] combine signature techniques with Paillier homomorphic encryption to produce a strategy suitable for the edge layer..

uses pseudonyms to provide device anonymity, and the devices are certified using the concept of a pseudonym certificate [14]. Data is aggregated using Paillier homomorphic encryption, with the intended application being fog-enhanced Internet of Things systems. provides a way to aggregate data for Internet of Things applications enabled by mobile edge computing while maintaining source integrity, authentication, and privacy [13].

Aggregation occurs in cloud and fog conditions as well. A strategy for using homomorphic encryption, for instance, to compile data from smart meters in a fog environment is presented in [16]. Homomorphic encryption is also used by Wang et al. [17] to gather information in the mist and move it to the cloud. An extra procedure that allows for the dynamic joining and leaving of terminal devices has been introduced by Shen et al. [18].

In cloud and fog situations, Intel SGX handles sensitive IoT data. Through the use of middleware, Gremaud et al. [19] demonstrate how to handle IoT data in the cloud while protecting privacy by hiding data from both the cloud provider and unauthorized parties. Silva et al. [20] take into consideration both homomorphic encryption and Intel SGX as two techniques for data security and privacy in their architecture for data aggregation in cloud computing. The architecture is used in a smart grid scenario. Lastly, it has been shown that using Intel SGX for data anonymization and aggregation permits privacy-aware data transmission, which is more practicable than homomorphic encryption.

The recommended solutions for cloud and fog environments' data aggregation, although focusing on a particular application context, despite many developments in the industry and the variety of methodologies used, do not manage heterogeneous data or data from multiple types of devices. Therefore, these techniques cannot be used to create many Internet of Things applications that allow for the dynamic addition and removal of devices from the data gathering network as well as the alteration and fusion of data from several sources [5].

V. RECOMMENDED DESIGN

There are now more options for processing on sensitive data thanks to the introduction of Trusted Execution Environments (TEE), particularly Intel SGX technology. In this research, we offer a novel method of IoT data aggregation that leverages Intel SGX's enclaved execution to safeguard user and data privacy. In order to offer safe channels of communication for the transfer of data from edge gateways or Internet of Things devices to the fog or cloud environment, we completely hide the data from the cloud provider—a potentially suspicious third party. An example of the architectural schema is shown in Fig. 3.

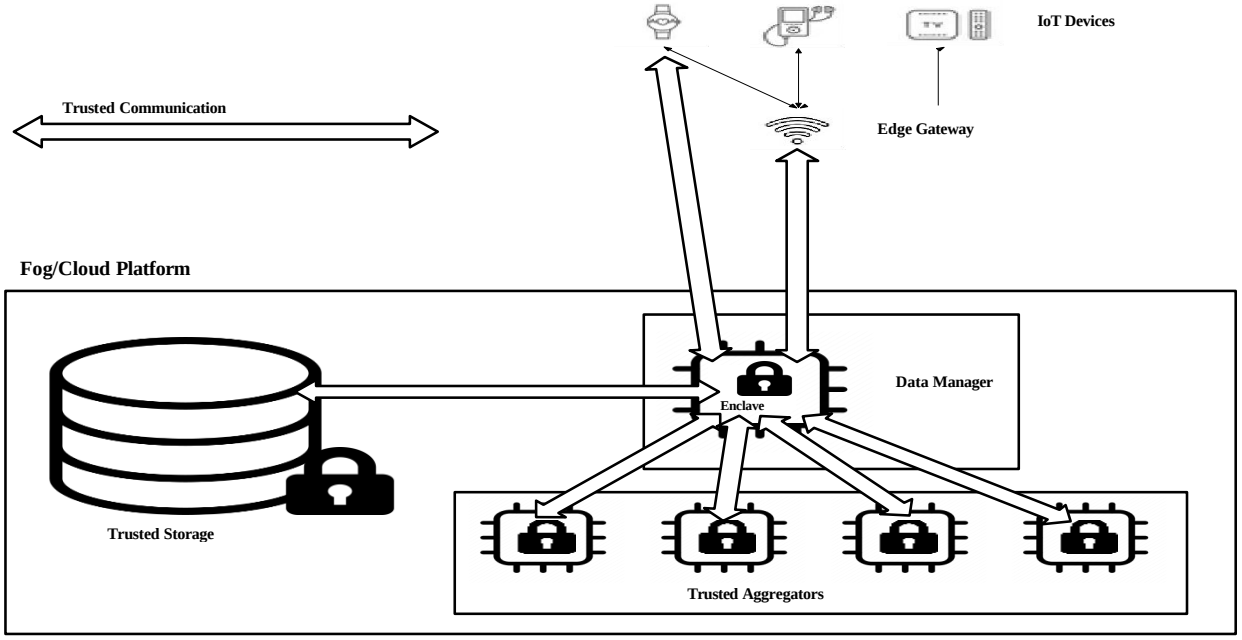


Figure 3 illustrates the suggested solution's architecture.

A data manager is required by the system to develop and oversee a number of enclaves, which we refer to as

trustworthy aggregators, to perform the aggregation functions. Sophisticated operations on encrypted data can be carried out at full processor speed with SGX enclaves, which lowers overhead and ensures data privacy [22]. One of Intel SGX's advantages is its ability to apply neural network and machine learning methods to sensitive and confidential data [23–26].

The data manager can arrange for the administration of several aggregators. An enclave inside the data manager called the enclave controller is also in charge of handling sensitive data. Once all the data from the edge gateway or Internet of Things devices has been collected, it selects which aggregator will perform the calculations. The enclave controller ensures consistent handling of all device requests and data.

All communication between the devices, the edge gateway, and the enclave controller is established over a reliable channel during the remote attestation process. The SGX enclave can securely communicate with external parties over the network and verify that it was successfully instantiated and hasn't changed thanks to remote attestation, which is covered in Section II.

Communications between the trusted aggregators and the enclave controller are also carried out through a safe path by utilizing the local attestation method. By going through this procedure, all of the information that they exchange is protected from the untrustworthy source. Moreover, the owners' confidentiality and privacy are protected because Within the enclaves, the data is processed and encrypted.

Trusted storage can be used to securely store all of the data that the data manager obtains, along with the outcomes of the computations that the trusted aggregators complete in a fog or cloud environment, guaranteeing data integrity and confidentiality. Intel SGX's sealing feature is the cornerstone of reliable storage (see Part II). To encrypt and store all requested data in the persistent storage, the enclave controller asks the CPU for a sealing key. As an alternative, trustworthy aggregators can store data directly in cloud or fog storage.

When heterogeneous data is combined and processed with little overhead utilizing Intel SGX's enclaved execution, it is more effective than other methods that have been reported in scholarly works, including homomorphic encryption [21].

The architecture shown in Figure 3 allows multi-level data aggregation and can be replicated in cloud and fog settings. This is due to the possibility that cloud providers have more processing ability to perform complex data analysis and summarization in addition to being able to combine data from many fog environments.

VI. CONCEPT PROOF

In order to evaluate the suggested architecture, Currently, a proof of concept is being created. Preliminary results indicate that employing remote attestation provides strong guarantees for secrecy and data integrity; both sides, however, need to account for the communication overhead [27, 28]. To confirm the authenticity of the enclave and find out if SGX has been activated on the platform it resides on, the Intel Attestation Service (IAS) must be utilized in conjunction with SGX attestation. In light of this, we will also assess the remote attestation process between the Internet of Things devices, edge gateway, and enclave controller using OPERA [29]. As an alternative, data could be transmitted via asymmetric encryption, which lowers the computational burden by requiring a key agreement.

Every trusted aggregator is housed inside an enclave, therefore allocating memory during enclave construction incurs a significant computational cost that influences the solution's response time. To address this problem, we will use enclave management strategies like pool [30] and enclave sharing [27], which lower the overhead associated with enclave initialization when interacting with trustworthy aggregators.

VII. FINAL COMMENTS AND NEXT WORK

This study provided a novel approach to protect user and data privacy in cloud and fog environments when aggregating data. Our system leverages the Trusted Execution Environment (TEE) to deploy sophisticated algorithms such as machine learning and to manage heterogeneous data supplied by several devices. TEE provides privacy and secrecy features as well.

While it's still early in the process, a Evidence of concept for the proposed plan is being created right now. A robust adversarial model will be considered while evaluating the application in real-world Internet of Things

environments. The benefits and drawbacks of the suggested solution would be highlighted by a security analysis. Additionally, performance studies will be conducted to quantify the overhead resulting from encrypted communications and Intel SGX.

VIII. REFERENCES

[1] "Systematic literature review on the use of trusted execution environments to protect cloud/fog-based Internet of Things applications," IEEE Access, vol. 9, pp. 80 953–80 969, 2021, D. C. G. Valadares, N. C. Will, J. Caminha, A. Perkusich, Mirko Barbara Perkusich, and K. C. Gorgonio.

[2] In Internet of Things From Hype to Reality: The Road to Digitization, A. Rayes and S. Salam discuss "Fog computing." 2019 Springer, pp. 155–180

[3] "Edge computing: Vision and challenges," IEEE Internet of Things Journal, vol. 3, no. 5, pp. 637–646, 2016, W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu.

[4] In the Proceedings of the International Conference on Communication and Signal Processing, C. Arivazhagan and V. Natarajan present an assessment of fog computing paradigms, difficulties, and prospects in the Internet of Things.

IEEE, 2020, Chennai, India, pp. 385–389.

[5] A lightweight privacy-preserving data aggregation technique for fog computing-enhanced Internet of Things, R. Lu, K. Heung, A. H. Lashkari, and A. A. Ghorbani, IEEE Access, vol. 5, pp. 3302–3312, 2017.

[6] In the Journal of Network and Computer Applications, volume 97, pages 23–34, 2017, Pourghebleh B. and Navimipour N. J., "Data aggregation mechanisms in the Internet of Things: A systematic review of the literature and recommendations for future research."

[7] "ESDRA: An efficient and secure distributed remote attestation scheme for IoT swarms," B. Kuang, A. Fu, S. Yu, G. Yang, M. Su, and Y. Zhang, IEEE Internet of Things Journal, vol. 6, no. 5, pp. 8372–8383, 2019.

[8] A survey of privacy-preserving techniques for heterogeneous data types, M. Cunha, R. Mendes, and J. P. Vilela, Computer Science Review, vol. 41, p. 100403, 2021.

[9] "Intel Software Guard Extensions in Internet of Things scenarios: A systematic mapping study," in Proceedings of the 8th International Conference on Future Internet of Things and Cloud, N. C. Will, D. C. G. Valadares, D. F. d. S. Santos, and A. Perkusich. IEEE, 2021, Rome, Italy, pp. 342–349.

[10] "Innovative instructions and software model for isolated execution," in Proceedings of the 2nd International Workshop on Hardware and Architectural Support for Security and Privacy, F. McKeen, I. Alexandrovich, A. Berenzon, C. V. Rozas, H. Shafi, V. Shanbhogue, and U. R. Savagaonkar. Israel's Tel-Aviv: ACM, 2013.

[11] The paper "Leveraging Intel SGX technology to protect security-sensitive applications" was presented at the 17th International Symposium on Network Computing and Applications by J. Sobchuk, S. O'Melia, D. Utin, and R. Khazan. (2018) IEEE, Cambridge, MA, USA, pp. 1–5.

[12] "Innovative technology for CPU based attestation and sealing," in Proceedings of the 2nd International Workshop on Hardware and Architectural Support for Security and Privacy, I. Anati, S. Gueron, S. P. Johnson, and V. R. Scarlata. Israel's Tel-Aviv: ACM, 2013.

[13] "Privacy preserving data aggregation scheme for mobile edge computing assisted IoT applications," by X. Li, S. Liu, F. Wu, S. Kumari, and J. J. P. C. Rodrigues, IEEE Internet of Things Journal, vol. 6, no. 3, pp. 4755–4763, 2019.

[14] "APPA: An anonymous and privacy preserving data aggregation scheme for fogenhanced IoT," Z. Guan, Y. Zhang, L. Wu, J. Wu, J. Li, Y. Ma, and J. Hu, Journal of Network and Computer Applications, vol. 125, pp. 82–92, 2019

[15] "LVPDA: A lightweight and verifiable privacy-preserving data aggregation scheme for edge-enabled IoT," by J. Zhang, Y. Zhao, J. Wu, and B. Chen, IEEE Internet of Things Journal, vol. 7, no. 5, pp. 4016–4027,

2020.

[16] "A secure data aggregation protocol for fog computing based smart grids," by F. Y. Okay and S. Ozdemir, was published in the Proceedings of the 12th International Conference on Compatibility, Power Electronics, and Power Engineering.

IEEE, Doha, Qatar, 2018, pp. 1–6

[17] In Future Generation Computer Systems, vol. 78, pp. 712–719, 2018, H. Wang, Z. Wang, and J. Domingo-Ferrer, "Anonymous and secure aggregation scheme in fog-based public cloud computing."

[18] "A privacy-preserving data aggregation scheme for dynamic groups in fog computing," Information Sciences, vol. 514, pp. 118–130, 2020, X. Shen, L. Zhu, C. Xu, K. Sharif, and R. Lu.

[19] "Privacy-preserving IoT cloud data processing using SGX," P. Gremaud, A. Durand, and J. Pasquier, Proceedings of the 9th International Conference on the Internet of Things. 2019; Bilbao, Spain: ACM.

[20] "Security and privacy aware data aggregation on cloud computing," L. V. Silva, P. Barbosa, R. Marinho, and A. Brito, Journal of Internet Services and Applications, vol. 9, no. 1, 2018.

[21] "Secure and privacy-aware data dissemination for cloud-based applications," in the Proceedings of the 10th International Conference on Utility and Cloud Computing, by L. Sampaio, F. Silva, A. Souza, A. Brito, and P. Felber. ACM, 2017, Austin, TX, USA, pp. 47–56.

[22] IRON: Functional encryption utilizing Intel SGX, B. Fisch, D. Vinayagamurthy, D. Boneh, and S. Gorbunov, Proceedings of the 24th Conference on Computer and Communications Security. ACM, 2017, Dallas, TX, USA, pp. 765–782

[23] In the Proceedings of the 25th USENIX Conference on Security Symposium, O. Ohrimenko, F. Schuster, C. Fournet, A. Mehta, S. Nowozin, K. Vaswani, and M. Costa discuss "Oblivious multi-party machine learning on trusted processors." USENIX Association, Austin, TX, USA, 2016, pp. 619–636

[24] Slalom: Quick, verifiable, and private neural network execution in trusted hardware by F. Tramer and D. Boneh, Proceedings of the 7th International Conference on Learning Representations. 2019; New Orleans, LA, USA: ICLR.

[25] SecureTF: A secure TensorFlow framework, D. L. Quoc, F. Gregor, S. Arnautov, R. Kunkel, P. Bhatotia, and C. Fetzer, Proceedings of the 21st International Middleware Conference. ACM, 2020; Delft, Netherlands, pp. 44–59.

[26] "PAIGE: Towards a hybrid-edge design for privacy-preserving intelligent personal assistants," by Y. Liang, D. O'Keeffe, and N. Sastry, was published in the Proceedings of the Third International Workshop on Edge Systems. ACM, 2020, Heraklion, Crete, pp. 55–60

[27] "Using a shared SGX enclave in the UNIX PAM authentication service," N. C. Will and C. A. Maziero, Proceedings of the 14th Annual IEEE International Systems Conference. IEEE, 2020, Montreal, QC, Canada, pp. 1–7.

[28] "Trusted inter-process communication using hardware enclaves," in Proceedings of the 15th Annual IEEE International Systems Conference, N. C. Will, T. Heinrich, A. B. Viescinski, and C. A. Maziero. IEEE, 2021; Vancouver, BC, Canada, pp. 1–7.

[29] In the Proceedings of the 26th Conference on Computer and Communications Security, G. Chen, Y. Zhang, and T.-H. Lai present "OPERA: Open remote attestation for Intel's secure enclaves." ACM Press, London, UK, 2019, pp. 2317–2331.

[30] "SGXPool: Improving the performance of enclave creation in the cloud," by D. Li, R. Lin, L. Tang, H. Liu, and Y. Tang, Transactions on Emerging Telecommunications Technologies, p. e3735, 2019.