Cloud Infrastructure With Confidential Computing

Kelan Morgan
Cloud Computing
South East Technological University
4th Year Software Deveolpment

***Abstract*—As cloud computing becomes more and more important to software infrastructure, ensuring the security of sensitive data while in use—not just at rest or in transit—has emerged as a critical concern. Confidential Computing is an emerging practice that addresses this challenge through the use of Trusted Execution Environments (TEEs), which offer hardware-based isolation for data and computation. Examples of such TEEs include Intel SGX [1], AMD SEV [2], and ARM TrustZone [3]. This paper explores the motivation behind Confidential Computing, provides a technical overview of its architecture, and examines real-world use cases in sectors such as finance and AI workloads. We assess its practical implications and performance-security trade-offs using case studies, and discuss the current limitations and future challenges in deploying Confidential Computing across cloud environments.**

I. Introduction

Cloud computing has revolutionized how organizations store, process, and manage data. However, as more sensitive data and worklods move over to public clouds which are easly accessed, there are increased concerns around data confidentiality, particularly during processing, have increased a lot. Traditional security measures such as encryption while it is stored and traveling across the network are becoming insufficient for protecting data in use — such as, data being processed in memory. This gap has driven the development of Confidential Computing, a important new part of cloud infrastructure aiming to ensure data remains protected throughout its entire lifecycle.

This report explores the architecture, existing and very impressive technologies, and practical implications of Confidential Computing in cloud environments. The goal is to provide a comprehensive overview accessible to those familiar with general cloud computing principles but not necessarily versed in secure hardware or cryptographic techniques.

This report contains the following:

1. Explaination of the concept of Confidential Computing and its enabling technologies.

2. Comparative analysis of major Confidential Computing platforms and cloud provider offerings.

3. Use cases and identification of limitations that need to be addressed for broader adoption.

II. Related Work

Confidential computing is a relatively new but growing field in cloud security. It builds on early work involving TEEs like Intel Software Guard Extensions (SGX), which provides per-process enclaves that isolate code and data from other applications and the OS [1]. In parallel, AMD's Secure Encrypted Virtualization (SEV) extends protection to full virtual machines with memory encryption and isolation at the hypervisor level [2]. ARM's TrustZone offers a system-level secure world used in mobile and embedded systems [3].

Some of the most important contributions where provided by academics such as Costan and Devadas' seminal paper "Intel SGX Explained" which provided a comprehensive analysis of SGX's strengths and limitations [4]. Arnautov et al. proposed SCONE, a secure container framework built on top of SGX, which integrates well with containerized applications while maintaining performance [5].

Security concerns still persist. Research such as Foreshadow has shown that SGX enclaves can be vulnerable to transient execution attacks like speculative side channels [6]. To mitigate these, additional hardware and microcode updates are required, but vulnerabilities remain a pressing concern for real-world deployment.

Recent industry initiatives such as the Confidential Containers (CoCo) project under the Cloud Native Computing Foundation aim to extend TEEs to Kubernetes workloads, enabling confidential container deployment across multiple cloud providers [7].

A. Advancements in Confidential Computing for Machine Learning

The integration of confidential computing with machine learning (ML) has garnered significant attention, particularly in scenarios involving sensitive data processing across distributed environments. A survey by Zobaed and Salehi explores the deployment of Trusted Execution Environments (TEEs) in edge-to-cloud infrastructures, emphasizing their role in securing ML workflows [8]. The study highlights the necessity for robust attestation mechanisms and addresses challenges related to performance overheads and the complexity of managing TEEs in heterogeneous systems.

III. System Architecture

Confidential computing architecture is designed to protect data in use by performing computations within hardware-based Trusted Execution Environments (TEEs). The core components of this architecture include TEEs, remote attestation services, and secure runtime environments.

A. Trusted Execution Environments (TEEs)

TEEs provide isolated environments that ensure the confidentiality and integrity of code and data during execution.

Examples include Intel Software Guard Extensions (SGX), AMD Secure Encrypted Virtualization (SEV), and ARM TrustZone. These technologies enable secure execution by isolating sensitive computations from the rest of the system.

B. Remote Attestation Services

Remote attestation is a process that verifies the integrity of the TEE before sensitive data is processed. It ensures that the code running inside the TEE has not been tampered with and is running in a secure environment. This is crucial for establishing trust in cloud and distributed systems. Tools like Gramine, Veraison, and Keylime facilitate remote attestation across various platforms [9].

C. Secure Runtime Environments

Secure runtime environments, such as Kata Containers, are lightweight virtual machines that provide an additional layer of isolation. In Kubernetes environments, these containers are launched as pods and wrapped in TEEs using frameworks like Confidential Containers (CoCo). This setup allows for secure deployment and execution of containerized applications [10].

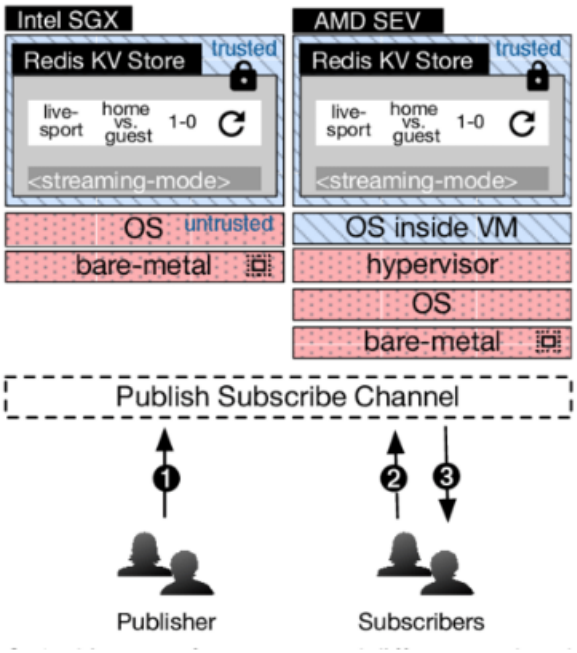

Fig. 1. Confidential Computing Architecture with Intel SGX and AMD SEV. Source: [11]

D. Attestation Process in Confidential Containers

In the context of confidential containers, the attestation process involves verifying the container runtime stack running within the confidential environment. This includes components like the kata-agent and supporting services [10]. The process ensures that both the virtual machine and the container runtime components are trustworthy before sensitive workloads are executed [9].

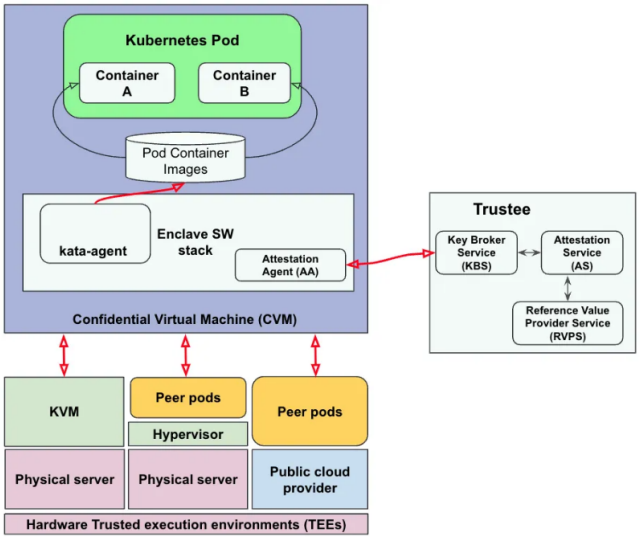

Fig. 2. Attestation Process in Confidential Containers. Source: [12]

E. Pod-Centric Approach in Confidential Containers

Confidential Containers adopt a pod-centric approach, balancing the Trusted Computing Base (TCB) size and resource sharing. This method places the workload pod and necessary helper processes within the enclave, while keeping other components like the hypervisor and control plane outside. This design enhances security while maintaining efficient resource utilization [13].

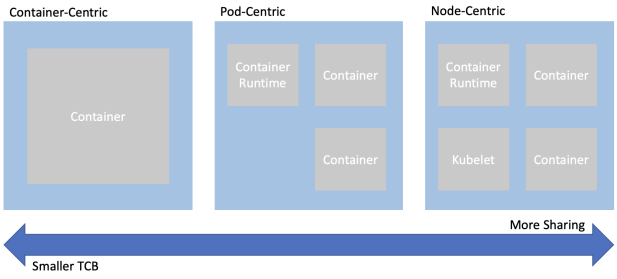

Fig. 3. Pod-Centric Architecture in Confidential Containers. Source: [13]

IV. Use Cases

Confidential computing is particularly beneficial in scenarios involving sensitive data, multi-party collaboration, or stringent regulatory requirements. By leveraging Trusted Execution Environments (TEEs), organizations can ensure data confidentiality and integrity during processing.

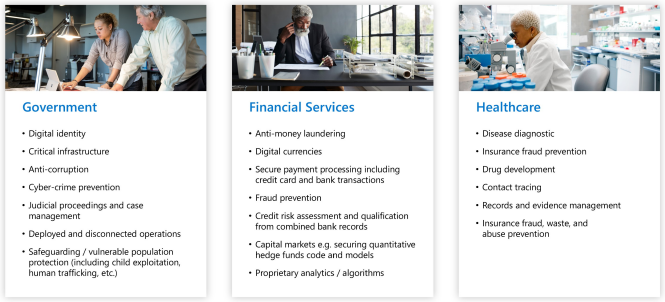

Fig. 4. Secure Multi-Party Computation in Banking. Source: [14]

A. Secure Multi-Party Computation Multiple institutions, such as banks or healthcare organizations, can collaboratively analyze data without exposing their individual datasets. For instance, in anti-money laundering (AML) efforts, banks can share encrypted transaction data within a secure enclave to detect fraudulent activities without revealing customer information. This approach enhances fraud detection while maintaining data privacy [15].

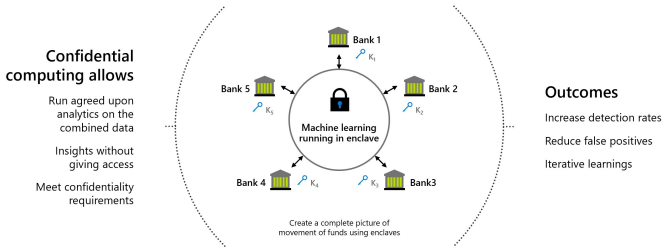

Fig. 5. Secure Multi-Party Computation in Banking. Source: [14]

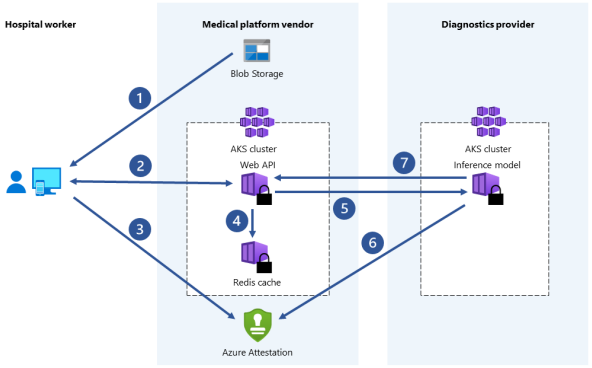

Fig. 6. confidential healthcare system. Source: [14]

B. Privacy-Preserving AI Machine learning models often require access to sensitive datasets for training or inference. Confidential computing enables models to operate on encrypted data within TEEs, ensuring that neither the data nor the model parameters are exposed. This is particularly valuable in healthcare, where patient data confidentiality is paramount. Platforms like BeeKeeperAI facilitate secure clinical algorithm development by providing a zero-trust environment for data stewards and AI developers [16].

C. Secure Edge Computing Edge devices operating in untrusted environments, such as retail stores or autonomous vehicles, benefit from on-device computation that guarantees data integrity and confidentiality. By embedding TEEs in edge hardware, organizations can ensure that sensitive data is processed securely, even in remote or physically insecure locations. This approach is crucial for applications requiring real-time data analysis and decision-making.

D. Intellectual Property Protection Organizations can safeguard proprietary algorithms and business logic by executing them within TEEs [17]. This ensures that intellectual property remains confidential, even when deployed in shared or public cloud environments. Industries such as finance and pharmaceuticals leverage this capability to protect sensitive computations from insider threats and external attacks [18].

E. Smart buildings utilize IoT devices to monitor occupancy, energy usage, and environmental conditions. These devices often collect sensitive data, such as CCTV footage or badge scans, which can be personally identifiable. Confidential computing enables processing this data within Trusted Execution Environments (TEEs), ensuring privacy and compliance with regulations like GDPR [14].

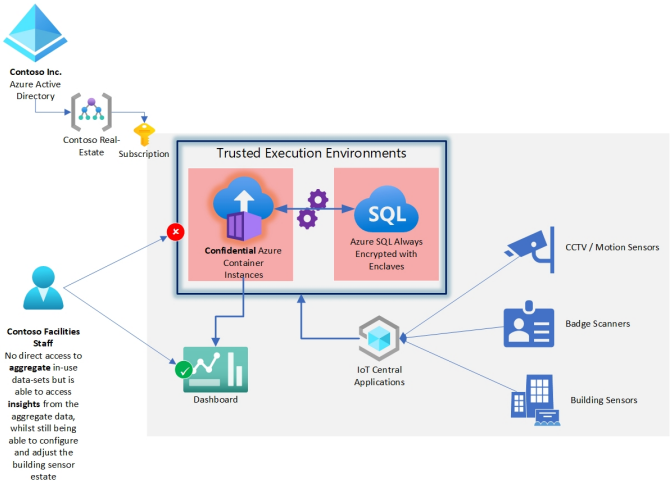

Fig. 7. Privacy-Preserving Data Processing in Smart Buildings. Source: [14]

F. Manufacturing processes often involve proprietary designs and techniques. When outsourcing production, there's a risk of intellectual property (IP) theft. By leveraging confidential computing, manufacturers can run sensitive computations within TEEs, ensuring that their IP remains protected even in untrusted environments [14].

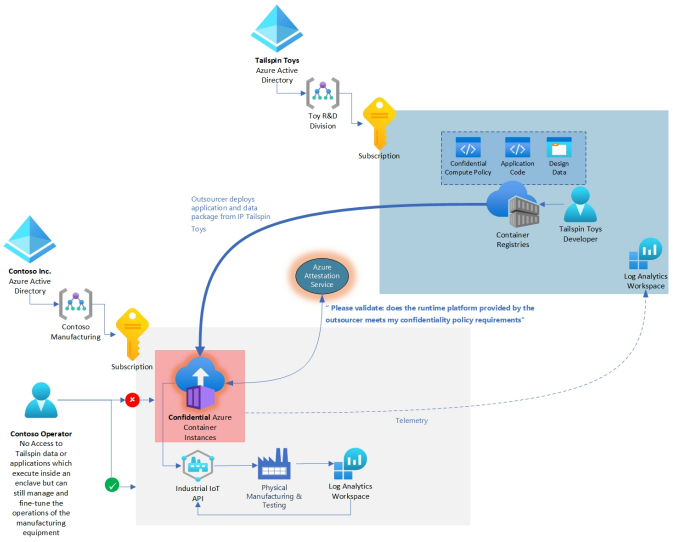

Fig. 8. Securing Manufacturing IP with Confidential Computing. Source: [14]

V. Case Studies

A. Accelerating Healthcare AI Development with BeeKeeperAI

The University of California, San Francisco's (UCSF) Center for Digital Health Innovation (CDHI) has developed BeeKeeperAI, a platform that leverages Intel® Software Guard Extensions (SGX) and Microsoft Azure Confidential Computing to facilitate the secure validation of clinical AI algorithms [19]. This approach enables algorithm developers to validate their models against sensitive healthcare data without exposing patient information or intellectual property. By utilizing secure enclaves, BeeKeeperAI ensures that both data and algorithms remain confidential throughout the validation process, significantly reducing the time and cost associated with developing clinical AI solutions [20].

B. Enhancing Financial Services Security through AMD and Google Cloud Collaboration

In a joint effort to bolster the security of confidential computing technologies, Google Cloud and AMD conducted a comprehensive audit of AMD's Secure Encrypted Virtualization-Secure Nested Paging (SEV-SNP) firmware. This collaboration involved Google's Project Zero and AMD's firmware team, aiming to identify and mitigate potential vulnerabilities in AMD's confidential computing stack. The audit led to the discovery and remediation of several security issues, reinforcing the integrity of AMD's SEV-SNP technology and setting a precedent for transparency and cooperation in enhancing cloud security [21].

VI. Limitations and Challenges

Despite its promise, confidential computing still faces key limitations:

- Hardware availability: Not all cloud instance types support TEEs, and they may incur additional costs.
- Performance penalties: Intel SGX enclaves are limited in memory (128MB usable by default), and entering/exiting enclaves adds latency [4]. AMD SEV provides larger memory scopes but also incurs I/O overhead.
- Complex development model: Developers must often use constrained SDKs and modify applications to run in enclaves. Debugging is also restricted due to security constraints.
- Side-channel vulnerabilities: Attacks such as Foreshadow and Spectre exploit microarchitectural features to leak secrets, even from secure enclaves [6].

These challenges must be addressed before Confidential Computing can see widespread production deployment.

VII. Conclusion

Confidential Computing is a promising development in cloud security, extending protection to data in use through hardware-based TEEs. While current technologies such as Intel SGX, AMD SEV, and ARM TrustZone offer solid foundations, there are still usability, performance, and security challenges to overcome.

We recommend further work in:

- Cross-platform APIs for enclave development
- Improved developer tooling and SDKs
- Research into secure multi-party computation on cloud-scale platforms
- Mitigation of side-channel attacks through hardware and compiler-level innovations

Confidential Computing is not a silver bullet, but as part of a larger defense-in-depth strategy, it offers strong guarantees for sensitive cloud workloads.

References

[1] Intel Corporation, "Software guard extensions (sgx)," 2020, https://software.intel.com/sgx.

[2] Advanced Micro Devices, Inc., "Secure encrypted virtualization (sev)," 2020, https://developer.amd.com/sev/.

[3] ARM Ltd., "Trustzone technology," 2020, https://developer.arm.com/technologies/trustzone.

[4] V. Costan and S. Devadas, "Intel sgx explained," in *IACR Cryptology ePrint Archive*, vol. 2016, 2016, p. 86. [Online]. Available: https://eprint.iacr.org/2016/086.pdf

[5] S. Arnautov *et al.*, "Scone: Secure computing with untrusted operating system," in *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*, 2016, pp. 689–703.

[6] J. V. Bulck *et al.*, "Foreshadow: Extracting the keys to the intel sgx kingdom with transient out-of-order execution," in *27th USENIX Security Symposium (USENIX Security 18)*, 2018, pp. 991–1008.

[7] Confidential Containers (CNCF), "Confidential containers project," 2023, https://confidential-containers.github.io.

[8] M. A. Salehi, "Confidential computing across edge-to-cloud for machine learning: A survey study," 2023, https://arxiv.org/abs/2307.16447.

[9] S. Kimmich, "What is remote attestation? enhancing data governance with confidential computing," 2024, https://confidentialcomputing.io/2024/10/02/what-is-remote-attestation-enhancing-data-governance-with-confidential-computing/.

[10] P. B. Ariel Adam, "Introducing confidential containers trustee: Attestation services solution overview and use cases," 2024, https://www.redhat.com/en/blog/introducing-confidential-containers-trustee-attestation-services-solution-overview-and-

[11] ResearchGate, "Architecture of our system and differences when deployed with intel sgx and amd sev-es," 2019, https://www.researchgate.net/figure/Architecture-of-our-system-and-differences-when-deployed-with-Intel-SGX-and-AMI fig3_330472504.

[12] R. Hat, "Introducing confidential containers trustee: Attestation services solution overview and use cases," 2023, https://www.redhat.com/en/blog/introducing-confidential-containers-trustee-attestation-services-solution-overview-and-use-cases.
[13] C. Containers, "Design overview - confidential containers," 2024, https://confidentialcontainers.org/docs/architecture/design-overview/.
[14] M. Azure, "Common azure confidential computing scenarios and use cases," 2025, https://learn.microsoft.com/en-us/azure/confidential-computing/use-cases-scenarios.
[15] ——, "Common azure confidential computing scenarios and use cases," 2025, accessed: 2025-05-18. [Online]. Available: https://learn.microsoft.com/en-us/azure/confidential-computing/use-cases-scenarios
[16] TechTarget, "Fostering health ai development with confidential computing," 2024, accessed: 2025-05-18. [Online]. Available: https://www.techtarget.com/healthtechanalytics/feature/Fostering-health-AI-development-with-confidential-computing
[17] M. Scapicchio, "What is confidential computing," 2024, https://www.ibm.com/think/topics/confidential-computing.
[18] A. Security, "Confidential computing use cases: Protecting sensitive data in a range of industries," 2023, accessed: 2025-05-18. [Online]. Available: https://www.anjuna.io/blog/confidential-computing-use-cases-protecting-sensitive-data-in-a-range-of-industries
[19] Fortanix, "Fortanix unlocks the power of confidential computing," 2021, https://www.fortanix.com/company/pr/2021/04/fortanix-unlocks-the-power-of-confidential-computing.
[20] Intel, "Intel sgx helps ucsf propel medical device innovations," 2020, https://download.intel.com/newsroom/archive/2025/en-us-2020-10-14-intel-sgx-helps-ucsf-propel-medical-device-innovations.pdf.
[21] L. H. Newman, "Amd gave google cloud rare access to its tech to hunt chip flaws," 2022, https://www.wired.com/story/google-cloud-amd-confidential-computing-security-audit/.