White Paper

Intel Security

Protecting Kubernetes Clusters in the Cloud with Confidential Computing and Intel® Xeon® Processors

Edgeless Systems Constellation uses Intel® Trust Domain Extensions to help secure Kubernetes

Authors

Dr. Benny Fuhry

Confidential Computing Architect, Intel Corporation

Dr. Felix Schuster

CEO, Edgeless Systems

Today, data is often encrypted at rest and in transit across the network, but not while in use by the processor and memory. Confidential Computing is an emerging industry initiative focused on securing data in use, without exposing it to the rest of the system. Security is often seen as code for containing, locking down, or deactivating data. In contrast, Confidential Computing is designed to unleash data and enable businesses to both transform and collaborate in ways that have been previously inaccessible.

Intel Corporation and **Edgeless Systems** are working to ensure that Confidential Computing extends to Kubernetes clusters, helping secure the flexibility of container deployment so data owners can better maximize their investment in information technology.

Confidential Computing

Addressing the increased need for comprehensive protection that also safeguards data in use, Intel spearheaded the drive to Confidential Computing, designed to protect data in use by performing computations in a hardware-based, attested Trusted Execution Environment (TEE).[1] For regulated or sensitive data, this level of continuous protection of workloads is critical and can provide organizations with the confidence to take advantage of the cloud's cost, scalability, and agility benefits.

The severity of cyber risks for organizations today has made risk prevention a central business issue. Studies have shown that cybercriminals can penetrate 93% of company networks,[2] and that the average cost of a data breach in the US is $9.44 million ($1.12 million globally).[3] In addition, nearly half of all data breaches occur in the cloud.[2] Testifying to the difficulty of dealing with intrusions, the average time to detect and contain a data breach is 277 days, but shortening that time to 200 days can save over $1 million per occurrence.[2]

Hardware-based TEEs enable organizations to more securely perform data operations in an isolated environment where application security and data confidentiality are protected, even if the environment is not trusted by the organization. In addition, TEEs provide attestation, a cryptographic proof that the desired processing has occurred inside a protected area, helping guarantee process integrity.

Making Confidential Computing Possible

Recognizing what was at stake as security risks increased, Intel conceptualized, developed, and released **Intel® Software Guard Extensions (Intel® SGX)** in 2015. This groundbreaking, hardware-based technology was a big step forward as it introduced the first highly performant data protection in-use technique combined with attestation (as illustrated in Figure 1). Intel SGX enables developers to better protect individual applications from other software and actors inside a TEE. This TEE provides data confidentiality, data integrity, and process integrity for applications. Combining Intel SGX with data protection at rest and in flight enables organizations to confidently run workloads everywhere, even if the environment is not (fully) trusted. Because security functions were built directly into advanced Intel® processors, sensitive data is processed at high speed.

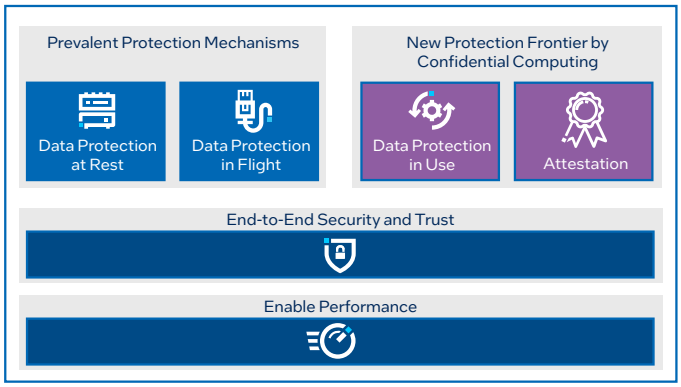

Figure 1. Security every step of the way.

In 2023, Intel has extended hardware-based security even further with the introduction of **Intel Trust Domain Extensions (Intel TDX)**. As shown in Figures 2 and 3, Intel TDX erects a protection perimeter, also known as a "Trust Boundary," around Virtual Machines (VMs)—the popular software emulations of computing devices that facilitate simplified creation and management of a wide variety of compute workloads. VMs protected by Intel TDX are called Trust Domains, and they can contain an Operating System such as Ubuntu or Microsoft Windows, or a minimal kernel, as well as multiple applications. Trust Domains are designed to prohibit anything from the hardware stack accessing code or data within the Trust Domains' boundaries, including the host OS, hypervisor, firmware, or system administrator. Multiple Trust Domains can run on the same server. They are protected from each other using a combination of different ephemeral keys for main memory encryption and access controls inside the CPU.

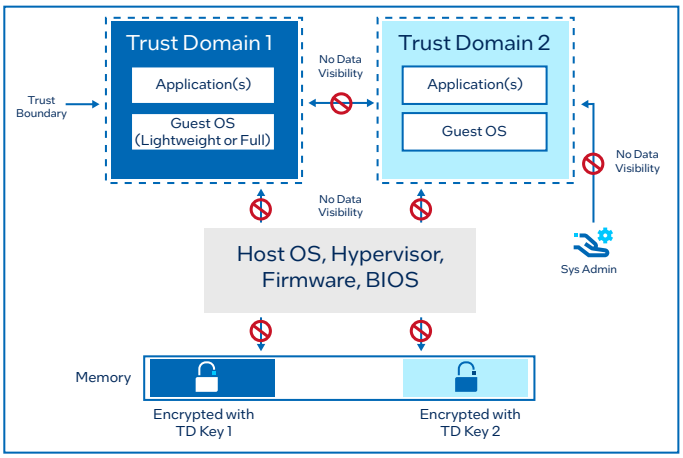

Figure 2. Intel TDX.

Thanks to hardware-enabled security built into Intel® Xeon® Scalable processors, Intel TDX can provide strong protection against both physical and software-based attacks:

- **Malicious Hypervisors**: In the wrong hands, hypervisors can be used to compromise a VM by intercepting and modifying sensitive data stored in VM memory, especially in cases where multiple VMs are running on the same system. Encrypting the entire memory area of a VM and securely handling VM operations (via the newly introduced Intel TDX Module) helps prevent unauthorized access to the VM's memory and helps ensure the VM remains secure even if the hypervisor is compromised.
- **Privileged Bad Actors**: Systems are designed to be administered by humans, so different levels of human access are built in. Anyone with malicious intent and privileged access (whether an insider or an outsider) can present a real security risk. Intel TDX is designed to prohibit the infrastructure administrator from accessing Trust Domains (i.e., VMs protected by Intel TDX). A customer administrator with legitimate access to the VM still has access by default, but such a remote access can be deactivated.
- **Rogue Guest Operating Systems**: In a virtualized environment, multiple VMs may reside on the same physical server, with each VM running a different guest OS. A compromised OS can attempt to access or modify data belonging to other VMs. Intel TDX helps prevent that because each Trust Domain's memory is encrypted with a dedicated key, enforcing strong, cryptographic isolation.
- **Physical Attacks**: In some cases, attackers can gain physical access to the platform. These incursions can involve the theft of platform memory (or even the platform itself), the installation of hardware-based malware, or the bypassing of other security measures. By encrypting the memory of Trust Domains, Intel TDX helps ensure that stolen memory cannot be decrypted by the attacker. Intel TDX also augments the Trust Domain's defense against limited attacks that seek physical access to the platform memory, such as offline DRAM analysis and cold-boot attacks. It also helps repel active attacks on DRAM interfaces, including capturing, modifying, relocating, splicing, and aliasing memory contents.
- **Side-channel Attacks**: Intel TDX encryption helps prevent side-channel attacks that seek to exploit vulnerabilities in platform hardware or software. Some of these attacks can be very sophisticated; for example, an attacker might try to exploit a timing vulnerability to access encryption keys.

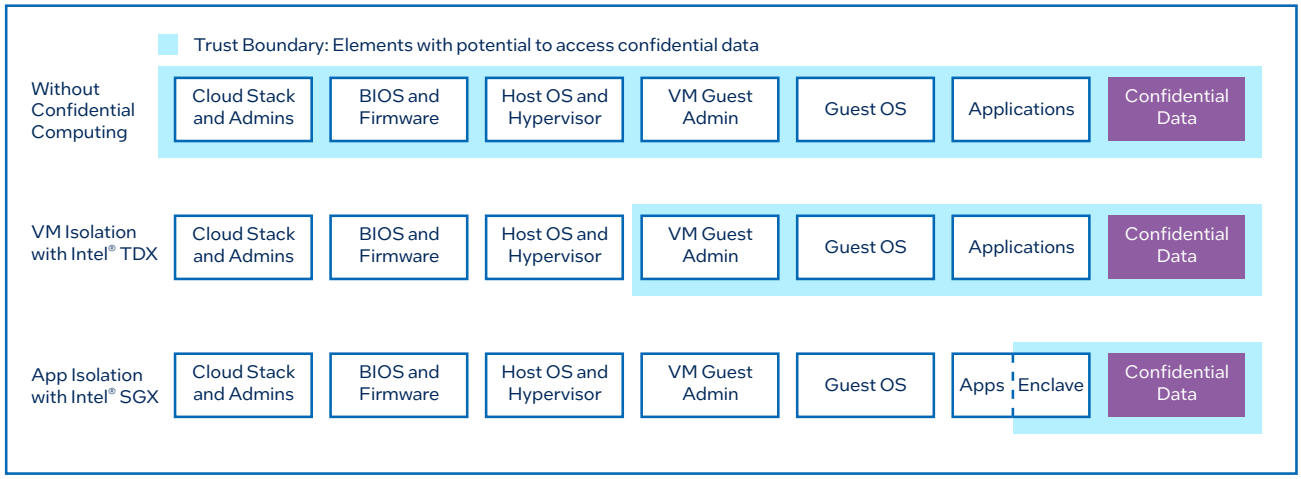

Figure 3. Trust Boundaries mapped to Confidential Computing with Intel TDX and Intel SGX.

Intel TDX Operational Highlights

Robust Trust Domains

Intel TDX provides robust Trust Domains featuring memory and CPU state confidentiality and integrity that help secure sensitive IP and workload data against many software- and hardware-based attacks. Software, firmware, devices, and even cloud platform operators can be excluded from the Trusted Computing Base (TCB). In addition, Intel TDX provides secure access to CPU instructions, debug operations, and separate security procedures, regardless of the cloud infrastructure used to deploy workloads.

Trust Boundaries

Intel TDX sharply discriminates between trusted and untrusted operations, with each Trust Domain completely surrounded by a Trust Boundary. Trust Boundaries define which operations can be run and which are prohibited, as shown in Figure 4.

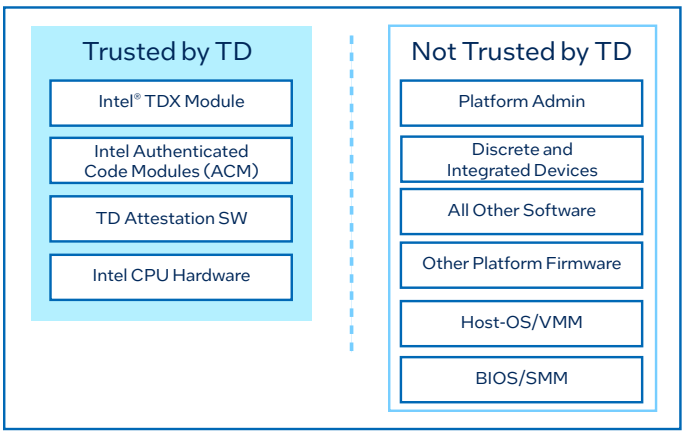

Figure 4. Trust Boundaries.

CPU Interaction

To help enforce the security policies for the Trust Domains, Intel introduced a new CPU mode on Intel Xeon Scalable processors called Secure Arbitration Mode (SEAM) to host an Intel-provided, digitally signed (but not encrypted) security-services module—the Intel TDX Module. This module is hosted in a reserved memory space identified by the SEAM range register (SEAMRR). The CPU allows access to this SEAMRR only for software executing inside that range; all other software and direct-memory accesses (DMAs) from devices attempted in this memory range are aborted. A CPU running in SEAM does not have memory-access privileges to other protected memory regions in the platform, including the System Management Mode (SMM) memory or memory protected by Intel SGX (Figure 5).

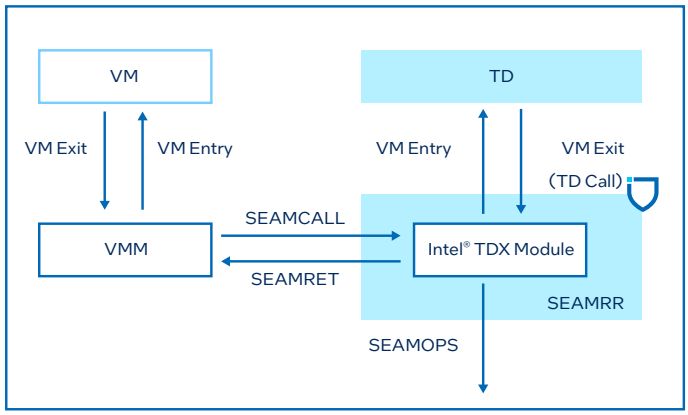

Figure 5. SEAM Module flow.

Attestation

Remote attestation provides stakeholders (data owners or service providers) with increased confidence that workloads are running inside a Trust Domain on a genuine Intel processor with Intel TDX functionality enabled. It also confirms that the system is running at the most current patch level.

With these capabilities, stakeholders can have increased confidence in the integrity of both the software and hardware components within the Intel TDX-protected environment.

The CPU provides a unique instruction that a Trust Domain can use to create an attestation report (i.e., a quote). To support attestation, Intel TDX uses a general-certification infrastructure built on the Intel® Software Guard Extensions Data Center Attestation Primitives (Intel® SGX DCAP) to help confirm that the Trust Domain in question has a certificate chain rooted to a valid, Intel-issued certificate (see Figure 6).

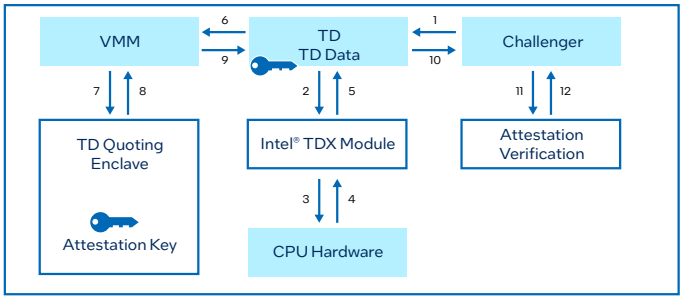

Figure 6. Attestation through Intel TDX.

Edgeless Systems

Edgeless Systems has been a leader in leveraging the capabilities of Intel SGX. They've created a series of powerful software packages that deliver new levels of flexible protection for data. Those products include:

- **EGo**, an open-source software development kit that enables organizations to develop their own confidential applications in the Go programming language. Developers can easily integrate Intel SGX features by adding a single line of code.

Edgeless Systems Constellation: Leveraging the power of Intel TDX

Constellation is a confidential Kubernetes distribution that has been updated to take advantage of Intel TDX to provide always-encrypted Kubernetes clusters, as illustrated in Figure 7. Constellation applies VM-based isolation from the infrastructure to entire Kubernetes clusters. In essence, Constellation provides Kubernetes administrators with highly trustworthy Kubernetes clusters on the public cloud; these can be used just like normal Kubernetes clusters.

Intel TDX enables developers to isolate entire VMs and encrypt them at runtime. But that alone isn't enough to deliver trustworthy Confidential Computing. To isolate an entire Kubernetes cluster from the infrastructure, a tool must protect the entire control plane, the network, and storage. Further, to be practical, that tool must be easy to use and must integrate smoothly with existing tools and workflows.

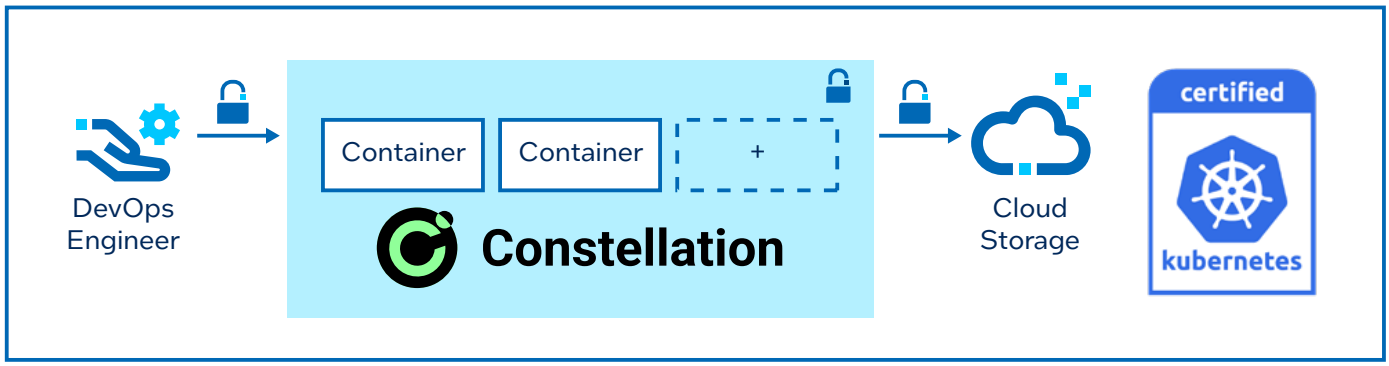

Figure 7. Isolating Kubernetes clusters.

- **EdgelessDB**, a full SQL database tailor-made to meet the requirements of Confidential Computing. It seamlessly integrates with customers' existing tools and workflows to help unlock the full potential of data while keeping that data secure.
- **MarbleRun**, an open-source control plane that simplifies the deployment, scaling, and verification of Intel SGX-based applications. It is designed to run on Kubernetes, either alongside a service mesh or as a standalone service.

Edgeless Systems Constellation addresses these requirements, extending its Confidential Computing capabilities beyond individual VMs to the entire cluster (including the control plane), as illustrated in Figure 8.

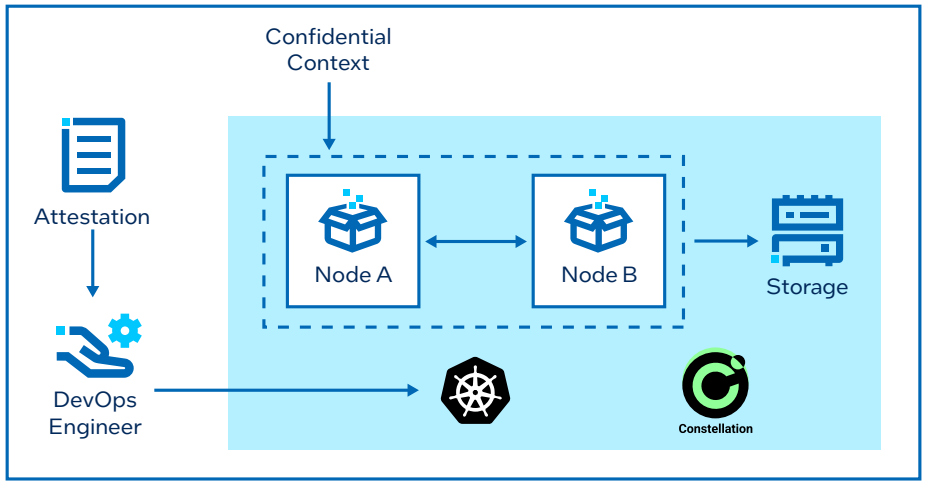

Figure 8. Protecting the entire cluster.

Constellation thereby helps ensure that all data in a cluster is always encrypted—at rest, in transit, and in use (i.e., during processing in memory). This way, Constellation aids in providing comprehensive protection against attacks from the infrastructure. These include attacks involving a compromised host operating system (e.g., through malware or rootkits) and malicious system administrators or datacenter employees. It also defends against hackers coming through a co-tenant.

Plus, Constellation provides true end-to-end attestation for a cluster—verifiable proof that all Kubernetes nodes are runtime-encrypted via Intel TDX and are running the intended system software. This enables customers to know exactly which underlying system software is running on the nodes, from the firmware to the operating system and the Kubernetes layer. This takes the concept of zero trust to the next level.

Constellation achieves this by measuring and recording each system software component that is loaded in a Confidential VM into the Runtime Measurement Registers (RTMRs) provided by Intel TDX. The "measured" components include firmware, bootloader, kernel, and a node's file system. The final values of RTMRs are included in the cryptographically signed attestation statements issued by the processors using Intel TDX (see Figure 9).

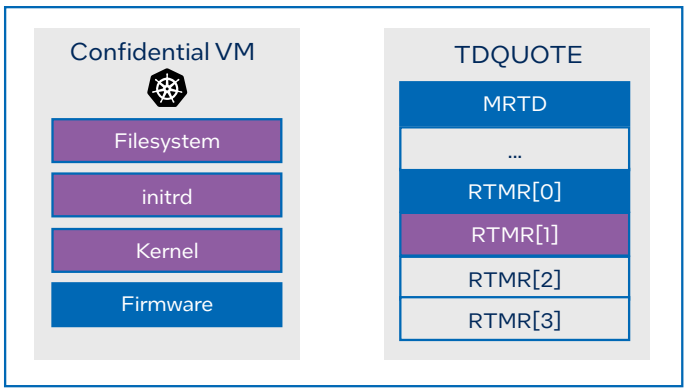

Figure 9. Usage of RTMRs to measure software parts.

Constellation's client-side frontend automatically verifies the attestation statement of the first node in a cluster. On success, the frontend sets up a secure connection to the first node and transfers the cluster's master secret, from which the cryptographic keys for network and storage encryption within the cluster are derived. The frontend also transfers a policy to the first node. This policy contains the acceptable RTMR values for nodes in the cluster and the cryptographic hashes of to-be-installed software components. For instance, these components include Constellation services for key management. The first node downloads and verifies these components and bootstraps the Kubernetes cluster. Crucially, the first node launches a service called Join Service that verifies new nodes based on the given policy. With this, the cluster can automatically scale securely without requiring user interaction.

Once a new node is verified, it receives cryptographic keys for network and storage encryption from the service. With these keys, the new node can interact with existing nodes in the cluster and can become a part of the cluster. Through this mechanism, Constellation ensures that all nodes in the cluster comply with the given policy and run precisely the intended system software and configuration. This process is illustrated in Figure 10.

Constellation configures the Container Network Interface (CNI) and the Container Storage Interface (CSI) in the cluster so that all network traffic between nodes is always encrypted and, in addition, all data written to local storage or cloud storage is automatically encrypted.

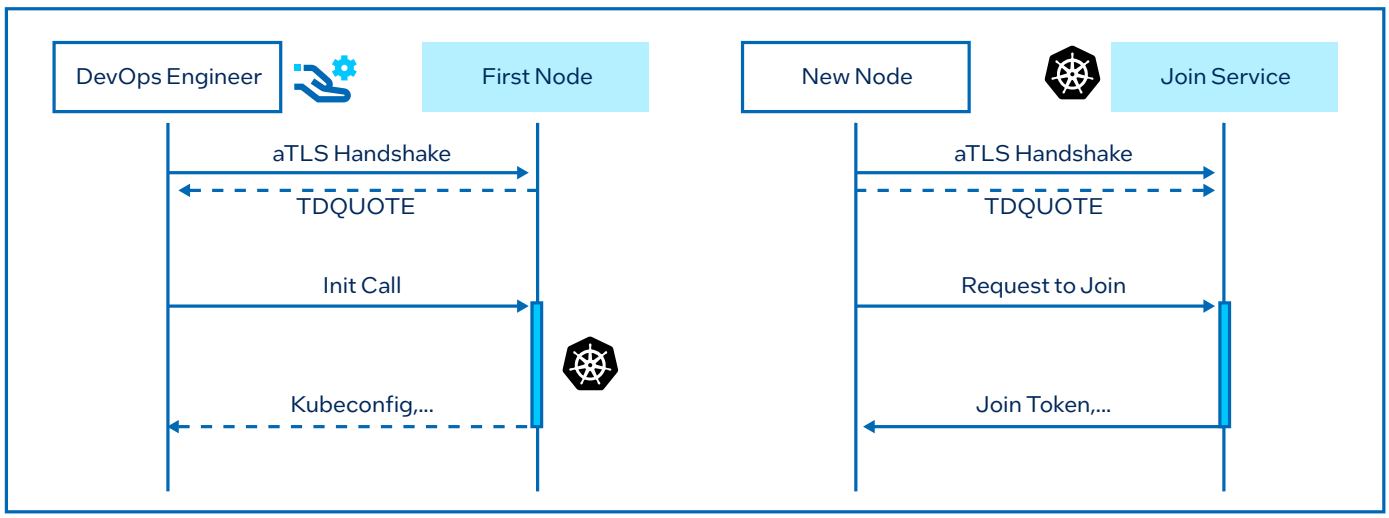

Figure 10. Setup of first Kubernetes node (left) and transitive trust used for other nodes (right).

Summary

In summary, Edgeless Systems Constellation delivers Confidential Computing at scale to customers by fully isolating and verifying Kubernetes clusters, enabling developers to quickly and easily run sensitive workloads on the cloud. Now every organization can use the public cloud with confidence.

Constellation benefits include:

- End-to-end encryption: designed to handle the most sensitive data and IP
- Performance and scale: high availability, autoscaling, and low overhead
- Easy deployment: Run with just a few commands on your CLI
- End-to-end verifiability: Cryptographic proof of security provisions
- Fast implementation: Run applications as they are, without code changes
- Open-source code

Also available is the Constellation Enterprise Edition, which comes with:

- Ready-to-use software packages
- Enterprise-grade features
- Long-term support releases
- Up to 24x7 support with guaranteed response times

Powering Confidential Computing

As we've seen, Intel TDX-based Confidential Computing can provide the foundation for the secure processing of Kubernetes workloads. With Intel Xeon Scalable processors, this secure processing can be delivered with high performance in agile, hybrid cloud data centers. Designed for multi-cloud workloads, they deliver high levels of performance, platform capabilities, and industry-leading workload acceleration, as well as hardware-based security. In addition, with trusted, hardware-enhanced data-service delivery and new I/O and connectivity technologies, these processors deliver improvements in I/O, memory, storage, and network technologies.

The Future: Staying on Guard

Computing technology, of course, is continually changing, constantly opening new opportunities and presenting new challenges. Nowhere is that more apparent than in Confidential Computing, where malefactors are always looking for loopholes to exploit and working on new attack strategies. Both Intel and Edgeless Systems understand that vigilance in staying on top of new security developments is necessary to maintain credibility as trusted providers of Confidential Computing technologies.

Unlike the never-ending challenges to security, the unrelenting focus on security that Intel and Edgeless Systems share is one thing that won't be changing any time soon.

Related Content

- **Edgeless Systems**
  www.edgeless.systems
- **Edgeless Systems Constellation**
  www.edgeless.systems/products/constellation
- **Intel Confidential Computing**
  https://intel.com/confidentialcomputing
- **Intel 4th Gen Xeon Scalable Processors**
  https://www.intel.com/content/www/us/en/products/docs/processors/xeon-accelerated/4th-gen-xeon-scalable-processors.html

NOTICES AND DISCLAIMERS

1. Common Terminology for Confidential Computing
https://confidentialcomputing.io/wp-content/uploads/sites/10/2023/03/Common-Terminology-for-Confidential-Computing.pdf
2. Alarming Cyber Statistics For Mid-Year 2022 That You Need To Know (forbes.com)
https://www.forbes.com/sites/chuckbrooks/2022/06/03/alarming-cyber-statistics-for-mid-year-2022-that-you-need-to-know/?sh=6d3075857864
3. Cost of a data breach 2022 | IBM
https://www.ibm.com/reports/data-breach

Intel TDX will be available on select 4th Gen Intel Xeon Scalable instances through four leading cloud providers. Previews have begun with select providers. Check with your provider for availability. Intel TDX becomes generally available with 5th Gen Intel Xeon Scalable processors. Intel technologies may require enabled hardware, software, or service activation. No product or component can be absolutely secure. Your costs and results may vary.

ACG6437TDX