

Powering Confidential Computing with AMI TruE™ using Intel® SGX

Contributors

Mourad Cherfaoui
Intel Corporation

Joseprabu Inbaraj
AMI

Abstract

With increased cloud adoption, securing workload in public cloud is extremely important. This white paper presents how AMI TruE™ and Intel® Security Libraries for Data Center (Intel® Secl-DC) provide data center operators with an infrastructure that allows tenants to leverage SGX remote attestation for securing their workloads, making them opaque to cloud administrators as well as intentional hackers.

Introduction

Confidential Computing (CC) has gained a lot of traction in recent years. This computing model protects sensitive data while it is being processed by a workload at runtime from other workloads running on the same platform. The industry has long addressed the problems of protecting sensitive data **at rest** and **in transit**. Data and transport channel encryption provide this protection. However, sensitive data still needs to be decrypted before it gets processed by a workload. The data is therefore exposed in RAM during processing, making it vulnerable. Data protection in use complements at rest and in-transit data protections, protecting data in all states.

Data in use is vulnerable to malware running on the same platform. If the workload is running in a data center with virtualization support, a malicious admin can just take a snapshot of a Virtual Machine to get all the sensitive data in RAM at that moment. Sensitive data leakage can also occur without malicious intent. It can happen because of mistakes on the part of the data center administrators.

CC addresses the gap of protecting data in use. One way to achieve this is by providing Trusted Execution Environments (TEEs) on the compute platform that helps ensure that data inside the TEE is protected from other workloads on the same platform¹. TEEs can also help ensure that code inside the TEE is protected against corruption. Through a process called **attestation**, TEEs can provide the ability for remote relying parties to verify that the workload is running in a genuine TEE. This allows advanced CC use cases. For example, after attesting that a workload is running in a TEE, a remote relying party can establish a secure channel with the workload with the assurance that sensitive data sent to the workload in the TEE is protected from other workloads on the same platform.

Hardware Security Modules (HSMs) and Trusted Platform Modules (TPMs) are examples of specialized TEEs that cannot be used to run generic workloads. Technologies that allow running generic workloads in TEEs have been introduced in recent years. These technologies typically rely on TEE support in the CPU rather than add-on hardware. This white paper focuses on the Intel® Software Guard Extension (Intel® SGX) technology. Intel® SGX implements the CC paradigm with support in the CPU.

This white paper introduces Intel® SGX and Intel® SGX attestation. It then describes how AMI TruE and Intel® Security Libraries for Data Center (Intel® Secl - DC) implement the infrastructure components needed by data center owners to support attestation of SGX workloads.

Table of Contents

Abstract	1
Introduction	1
Intel Software Guard Extension	2
Intel® SGX Attestation	2
Intel Security Libraries (Intel Secl-DC)	5
AMI TruE	6
Summary	8
References	8

Intel Software Guard Extension

Intel® SGX provides a set of Intel CPU instructions that allows the creation and management of TEEs called Intel® SGX **enclaves**. An Intel® SGX enclave is a region of memory containing data and code that is encrypted by a cryptographic key generated in the CPU package. Code and data in an enclave by design are only decrypted inside the CPU package (figure 1). The net result is that code and data in an enclave cannot be leaked to any software on the server, including privileged software like the OS kernel, the Virtual Machine Manager (VMM), the BIOS and the firmware (figure 2).

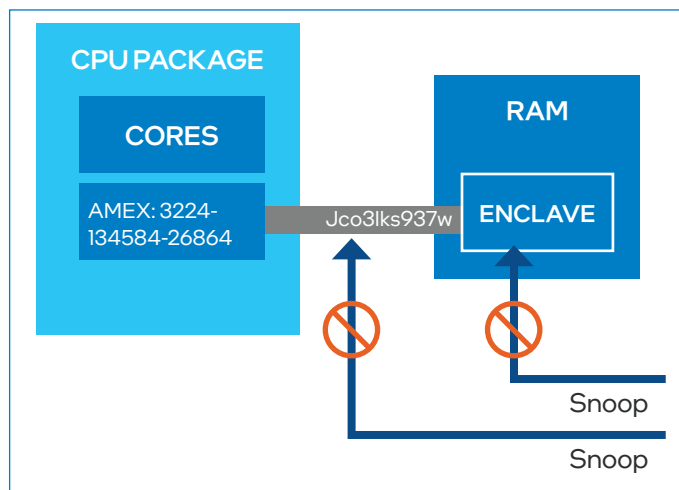


Figure 1. Data and code are protected in Intel® SGX enclaves

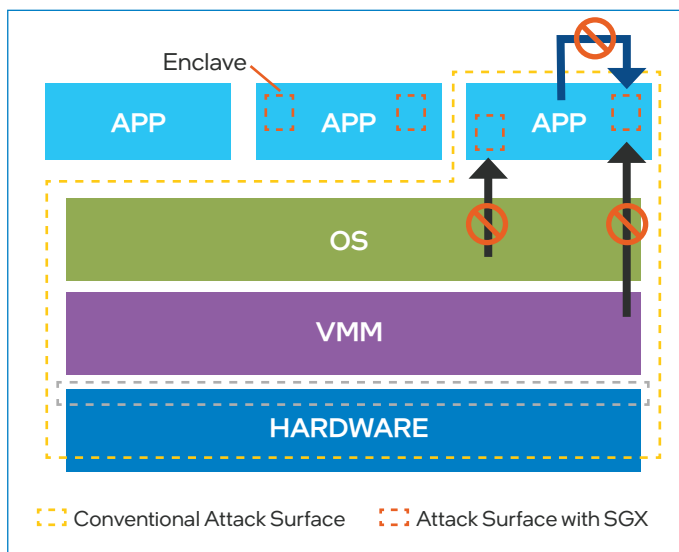


Figure 2. Contents of Intel® SGX Enclaves are not disclosed to any software

Intel® SGX enclaves are deployed as shared libraries. The code and data segments in a library are not encrypted, but they get encrypted by a CPU-generated key when the library is loaded. It is the responsibility of the application developer to identify the sensitive part of the application and define it as an Intel® SGX enclave library using the Intel® SGX SDK. The enclave part of the application is usually called the **trusted** part of the application, whereas the rest of the application is called the **untrusted** part. A good Intel® SGX application design must assume that the untrusted part of the application could be compromised. The untrusted and trusted parts of the application communicate by using dedicated Intel® SGX CPU instructions to maintain the segregation between the two domains. **eCalls** are used for communications from the untrusted part to the trusted part. **oCalls** are used in the reverse direction. Other than the communication protocol, from the untrusted part of the application, the Intel® SGX enclave (or trusted part) behaves exactly like any other shared library. In particular, the internal state of the enclave is not persisted when the enclave library is unloaded by the host application. There are instances however where we do want to persist some of the internal variables. Intel® SGX uses a mechanism called sealing to persist internal data beyond the time the enclave is loaded in RAM. Using sealing, Intel® SGX encrypts data using a CPU-generated key and stores it as an encrypted data blob on disk. There are two types of sealing: one that ties the data blob to the enclave's fingerprint (or **MREnclave**) and another one that ties it to the author of the enclave (or **MRSigner**). The type of sealing is defined by the enclave developer. The data blob can be unsealed during next runs of the enclave. The Intel® SGX hardware ensures that the data blob is only unsealed by an enclave that has the same MREnclave or the same MRSigner depending on the type of sealing used by the enclave's developer.

To support Intel® SGX workloads, the data center owner needs to offer platforms with Intel® SGX-enabled CPUs and the required software stack (Intel® SGX driver, Intel® SGX-aware VMM, Intel® SGX Platform Software or **PSW**). Intel® SGX has been supported on client and E3 server platforms. With the 3rd Gen Intel® Xeon® Scalable processors, Intel® SGX is also supported on E5 server platforms.

Intel® SGX Attestation

Intel® SGX attestation allows a remote relying party to verify that an Intel® SGX enclave is genuine. To support attestation at runtime, specific actions need to be taken by the enclave's developer, the data center owner, and the remote relying party.

Attestation support starts at development time. The enclave developer uses the Intel® SGX SDK to build and generate enclave attributes such as the enclave signer (**MRSigner**), the measurement (**MREnclave**), and the **Product Id**. MREnclave is a fingerprint of the enclave code and data segments. The MREnclave and the other attributes need to be communicated to remote relying parties that need to attest the enclave at runtime. Typically, the enclave source code is made public and remote relying parties can derive these attributes themselves so that they don't need to rely on the enclave developer to communicate the correct ones.

At runtime, the host application requests the Intel® SGX enclave to generate a hash-based message authentication code (HMAC)-signed **SGX report**, which the enclave does using the **ereport** CPU instruction. The report contains the MREnclave computed by ereport, the MRSigner, the Product Id and the other attributes, which ereport extracts from the enclave. The report can also contain data that the enclave wants to share with remote relying parties. This data is included in the **userData** field of the report. The **userData** field has a limited size so it usually contains a hash of the data instead. The actual data is communicated to remote relying parties separately. The remote relying parties can validate that the data comes from a genuine enclave by comparing its hash to the hash contained in **userData**.

The HMAC key used to sign the report is only accessible to the CPU and other enclaves on the same platform. Any enclave on the platform can therefore verify the report. This is called **local attestation**. Next, a **signed SGX quote** is generated using two PSW enclaves: the Quoting Enclave (QE) and the Provisioning Certification Enclave (PCE). The SGX quote contains the enclave's report and the platform Trusted Computing Base (TCB) information. The TCB information indicates the platform hardware and software security patch levels. The SGX quote is signed by the QE using its **Attestation Key** after validating the report's HMAC signature. The Attestation Key certificate is in turn signed by the PCE key. The PCE certificate is called the Provisioning Certification Key certificate (**PCK certificate**). The PCK certificate is issued by Intel during the platform SGX provisioning by the data center owner. The provisioning flow by the data center owner is detailed later in this white paper. Figure 3 shows the PSW enclaves and keys/certificates used to generate and sign an SGX quote.

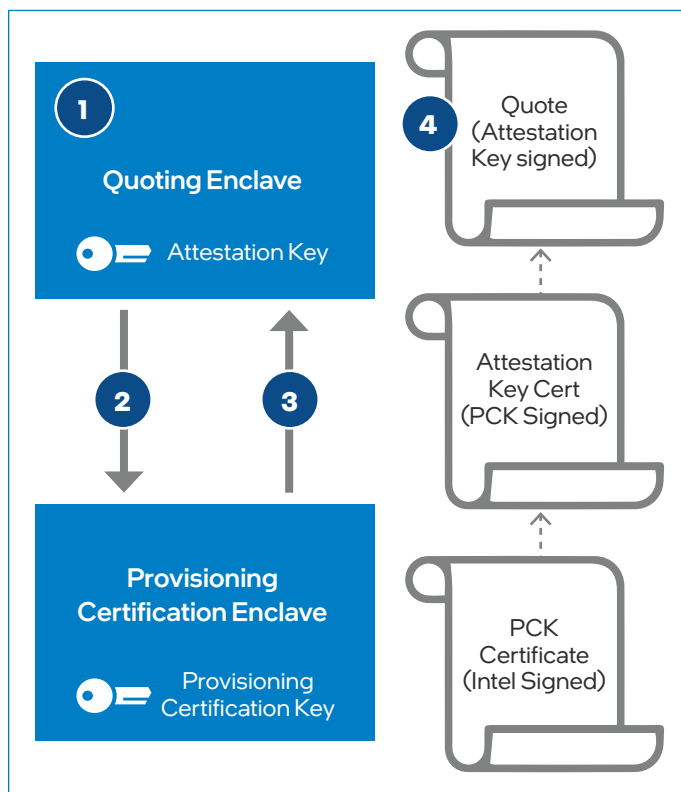


Figure 3. SGX Quote Signing Flow

Remote relying parties that need to attest an enclave receive an **extended SGX quote** comprised of the signed SGX quote, the signing certificate chain with the PCK certificate at the top of the chain and data returned by the enclave. The data is typically an enclave-generated public key. The remote relying party can check that the SGX quote is signed by a certificate that chains up to the PCK certificate, which is generated by Intel. The remote relying party ensures that enclave-generated data (such as a public key) is indeed generated in the enclave by checking that its hash matches the **userData** contained in the SGX quote². The SGX quote verification can also tell if the quote was generated on a platform that has the latest TCB. Figure 4 shows the attestation flow.

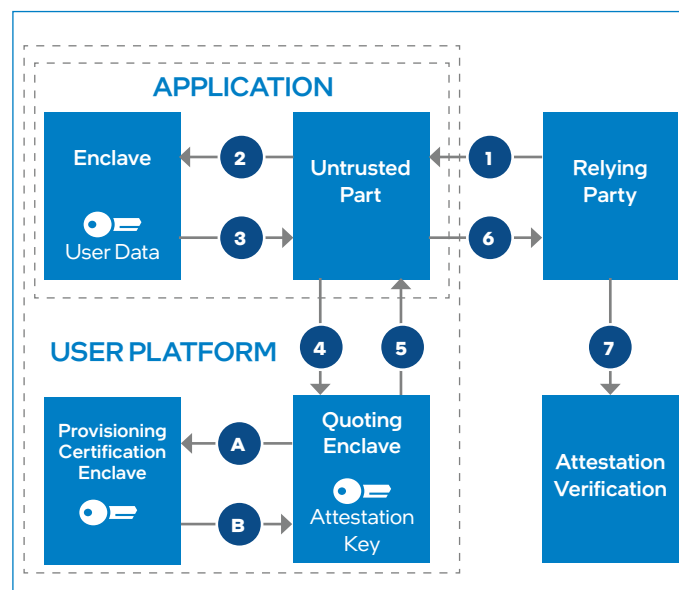


Figure 4. Attestation Flow

After verifying the SGX quote, the remote relying party may generate a Symmetric Wrapping Key (SWK), wrap it with the public key of the enclave (if it was returned as enclave data) and send it to the enclave. Only the enclave can unwrap the SWK. The remote relying party and the enclave can then exchange secrets wrapped with the SWK.

The SGX quote verification by remote relying parties can be policy-based. Beside verifying that the quote has a valid signature, the remote relying party may decide to accept the quote based on the TCB status and specific values of the quote attributes such as MRSigner and Product id.

As indicated before, the provisioning of the PCK certificates of the platforms is the responsibility of the data center owner. This process involves extracting SGX platform data from the platform. This data is then sent to the Intel Provisioning Certification Service (PCS), which generates the PCK certificates of the platform. PCS generates one PCK certificate per TCB level of the platform model. PCS is an Intel service available on the public Internet. This process uses an SGX Caching Service (SCS) proxy in the data center. The platform data is sent to SCS, which then contacts PCS. The certificates are cached in SCS and retrieved by SGX workloads when they need to generate an SGX quote. When it requests the platform PCK certificate from SCS, the SGX

workload extracts and provides the TCB level information in the request to SCS so that SCS returns the PCK certificate that corresponds to that TCB level. By fetching the PCK certificates for all the possible TCB levels of the platform in one call, SCS avoids repeated calls to PCS every time the TCB of the platform changes. The TCB changes when the platform firmware is upgraded.

As we saw, attestation involves many services and libraries for SGX platform provisioning, quote generation and verification. These libraries and services are made available by Intel to data center owners and remote relying parties that need to attest SGX enclaves. These libraries and services are collectively called Intel SGX Data Center Attestation Primitives (**DCAP**). The main primitives of DCAP are:

- The intel **SGX PCK Certificate ID Retrieval tool**: this tool extracts the SGX platform information that PCS uses to generate the PCK certificates of the platform.

- The Intel Quote Generation Library (**QGL**) and Quote Provider Library (**QPL**): these libraries are used to generate SGX quotes and to determine the TCB of the SGX workload. QGL uses the PSW QE and PCE enclaves.
- The Intel Quote Verification Library (**QVL**): this C library is used by remote relying parties to verify SGX quotes. **QVE** is a n implementation of QVL in an SGX enclave.
- The Provisioning Certification Caching Service (**PCCS**): this is a reference implementation of the SCS.

Figure 5 shows the DCAP components and how they are used during platform provisioning and at runtime.

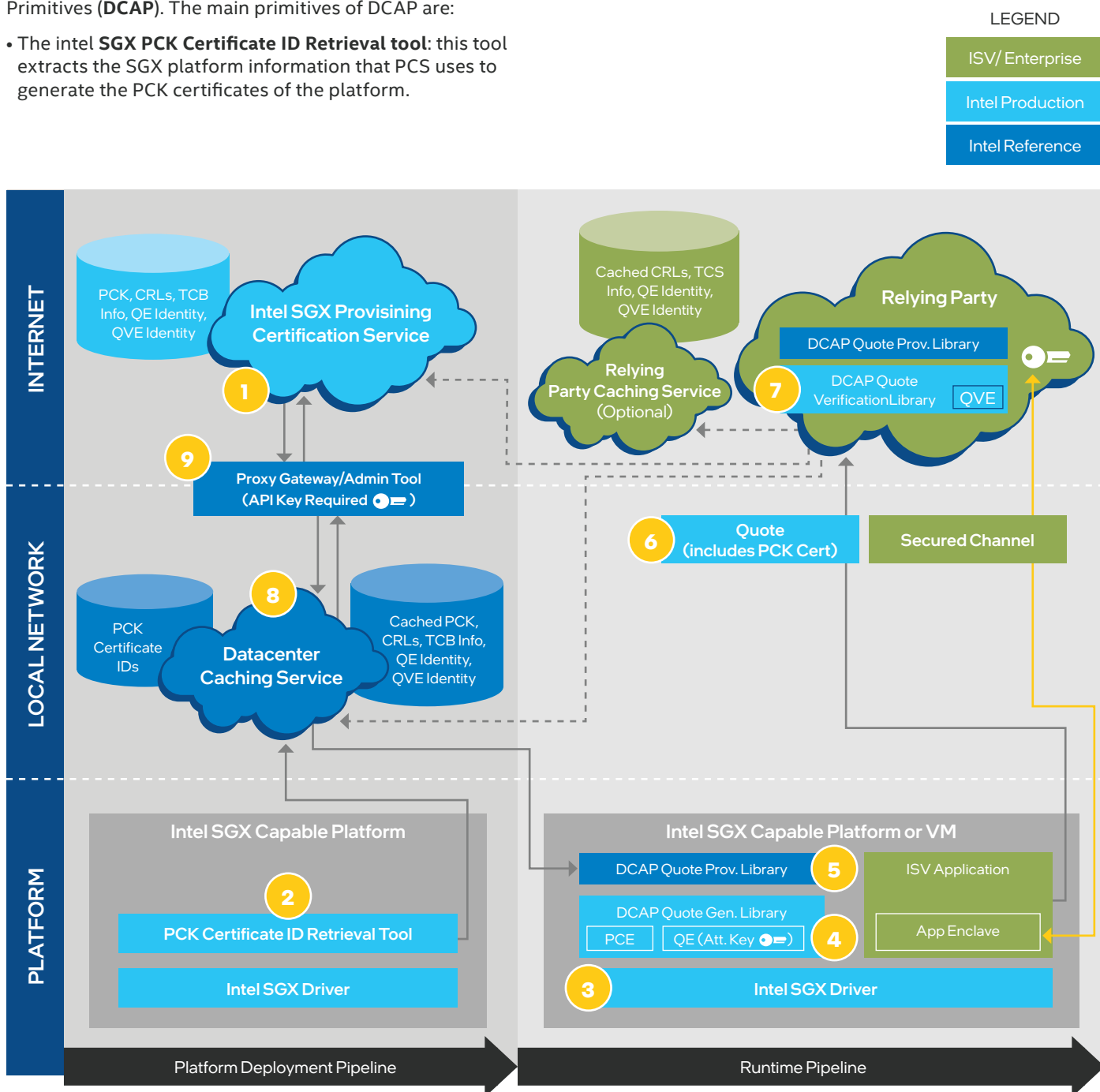


Figure 5. DCAP Components and Usage Flows

DCAP requires the data center owner and relying remote parties to build supporting infrastructures for authentication and authorization and manage the deployment of the primitives. Intel SecL-DC for Data Centers provide a turnkey solution built on top of DCAP that data center owners and remote relying parties can easily deploy to manage the SGX attestation.

Intel® SecL-DC

The Intel Secl-DC **SGX Attestation Infrastructure** pulls all the DCAP components together and adds other infrastructure services for authentication and authorization to provide a complete solution for supporting SGX workloads and SGX attestation in data centers and in remote relying parties' environments. Authentication and authorization in Intel Secl-DC are provided by the Certificate Management Service (**CMS**) and the Authentication and Authorization Service (**AAS**).

Figure 6 shows the SGX attestation infrastructure components and their interactions.

The main differences with DCAP are:

- The **Intel SecL-DC SGX Agent**: a wrapper of the SGX PCK Certificate ID Retrieval tool. It integrates with Intel SecL-DC CMS and AAS to add authentication and authorization

during the interaction with the Intel SecL-DC SGX Caching Service (SCS). Intel SecL-DC SGX Agent is a daemon that automatically updates SCS when needed.

- SCS: a Golang implementation of the PCCS reference implementation with a Postgres database backend for storing PCK certificates and platform collateral. SCS also integrates with CMS and AAS.
- The Intel SGX Quote Verification Service (**SQVS**): provides a REST API based implementation of the QVL C library. This allows relying parties written in any language to verify SGX quotes. SQVS also integrates with CMS and AAS. SQVS simplifies the policy decision making by remote relying parties.
- The solution includes a Secure Key Caching (**SKC**) use case. SKC protects keys in use inside SGX enclaves and leverages SGX attestation for provisioning keys into enclaves from remote Key Management Systems (KMS).
- The solution includes a sample application to demonstrate how to develop SGX workloads and how to use the SGX attestation infrastructure.

Additionally, the Intel SecL-DC SGX Attestation Infrastructure supports integration with data center management software such as orchestrators (OpenStack, Kubernetes) and **AMI TruE**.

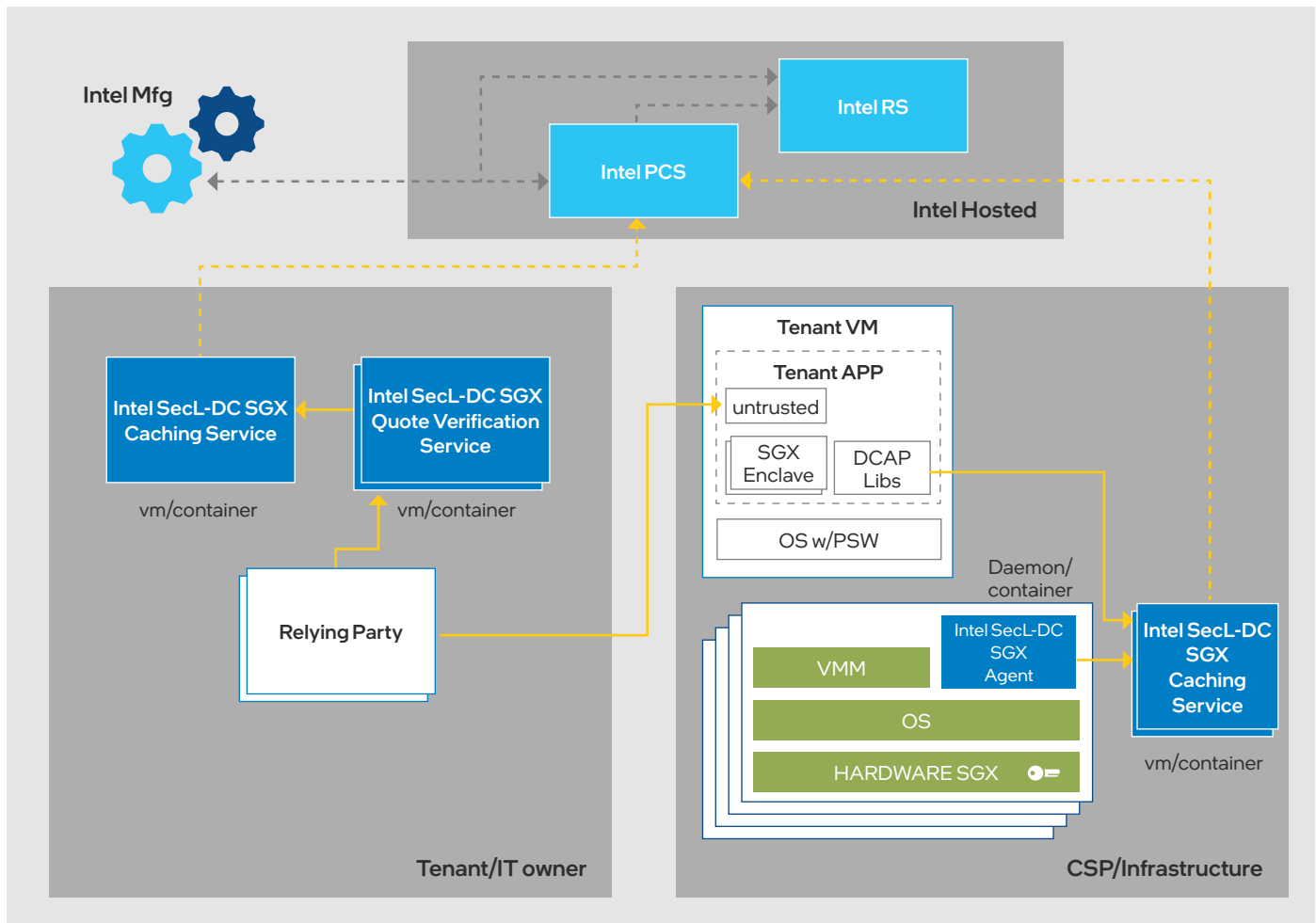


Figure 6. SGX Attestation Infrastructure in Intel SecL-DC

AMI TruE

AMI TruE provides holistic data center security solution using Intel® Security Technologies and Intel® Security Libraries for Data Centers to provide a Trusted Environment for cloud execution.

Leveraging Intel® SGX technology, AMI TruE enables confidential computing, eases deployment of workload attestation and secures application keys without compromising confidentiality. It is scalable, extensible and is built for cloud-to-edge applications. AMI TruE establishes and tracks the servers' trusted compute status in the data center, complies with data sovereignty regulations, runs sensitive workloads on trusted servers and provides remediation measures for untrusted platforms.

Deploying AMI TruE automatically installs required Intel® SecL-DC components including SGX Quote Verification Service (SQVS) and SGX Caching Service (SCS). Once deployed and configured, AMI TruE starts discovering servers and collects detailed asset information of all discovered servers. Following are the key platform and workload security features offered by AMI TruE:

- Discover manageable platforms.
- Identify platforms that have security features including TPM and Intel SGX.
- Remote provisioning of trust agents and SGX agents

- Platform trust management
- Platform integrity assurance
- Data sovereignty
- Monitor trust status of all TPM enabled platforms.
- Send email alerts when the trust status of a monitored platform is compromised.
- Enable remote remediation of untrusted platform, from updating BIOS/BMC firmware to reinstalling the operating system to rebooting the platforms.
- Manage secure workloads on SGX enabled platforms.
- Provision PCK certificates for SGX enabled platforms and SGX collateral.
- Label nodes in a Kubernetes cluster with trust status and SGX support
- Support for Secure Key Caching (SKC) in SGX enclaves.
- A sample application demonstrating how to develop SGX workloads and how to use the SGX attestation infrastructure.

To enable confidential computing, AMI TruE needs to be installed both in the CSP environment and in the tenant environment. The image below shows AMI TruE presenting a discovered SGX enabled host platform.

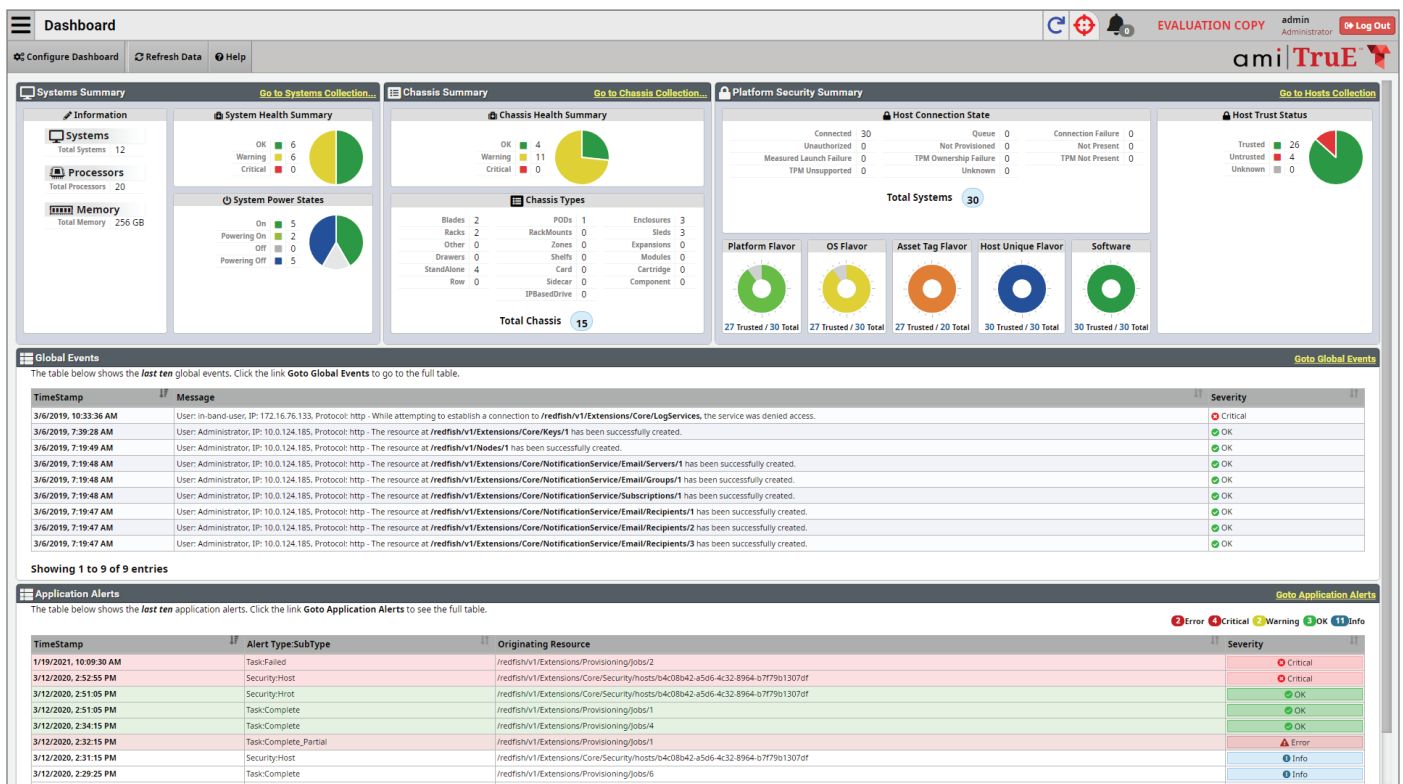


Figure 7. AMI TruE Dashboard showing health and security status of managed infrastructure

AMI TruE deployed in data centers allow administrators to add tenants who need SGX enabled hosts for launching secure workloads. Tenants can use AMI TruE to verify if an allocated host supports SGX to gain confidence that their workloads can be protected at runtime.

Once the assigned host is verified to have SGX support, the tenant creates workload on a host that has SGX enabled and, if desired, the latest TCB. This allows Kubernetes to verify and schedule SGX workloads on hosts that have SGX enabled.

The AMI TruE instance that is running in an enterprise environment provides two views in its web interface — host view and workload view. Host view interacts with Kubernetes (K8S) and shows all hosts assigned to this tenant. Workload view presents all running workloads managed by this tenant.

The screen shot in figure 8 shows host view which presents complete inventory information of SGX enabled hosts, including the list of workloads running in each host.

All workloads that are launched on SGX enabled host platforms are presented in a web interface as shown in Figure 9. View tenant details of a host

AMI TruE supports Kubernetes orchestration for launching workloads in a secure enclave. Future versions will support OpenStack and other orchestration solutions for secure workload launch. Every feature offered by AMI TruE is supported through REST API as well, enabling easy integration of AMI TruE with other data center management or orchestration solutions.

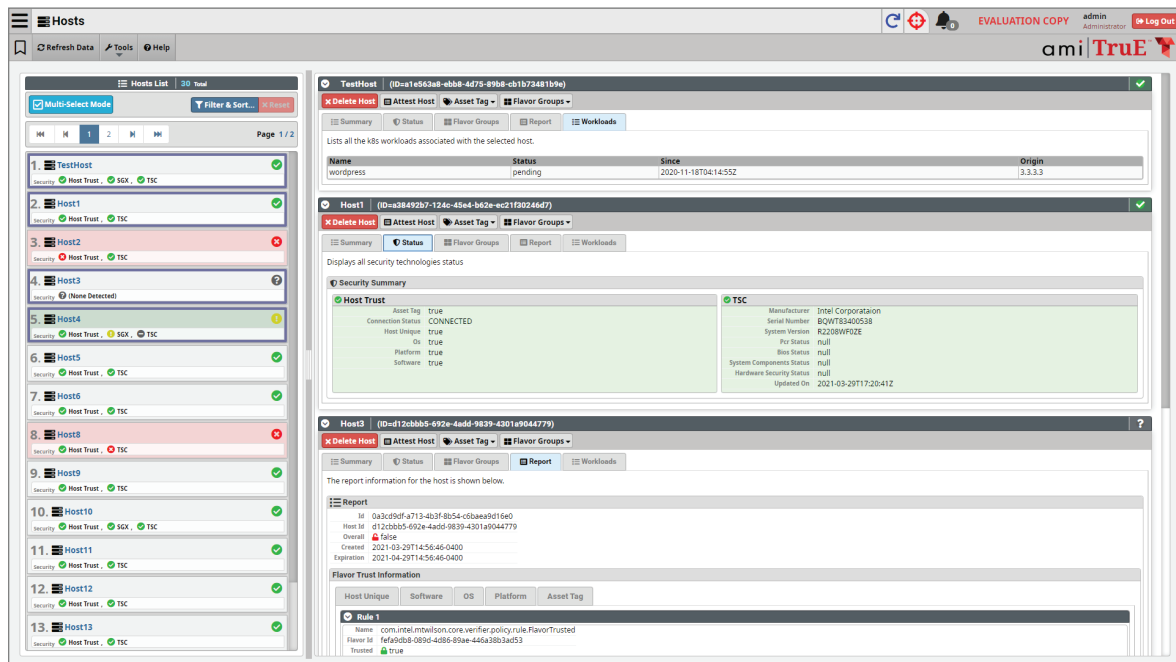


Figure 8. Manage hosts enabled with SGX

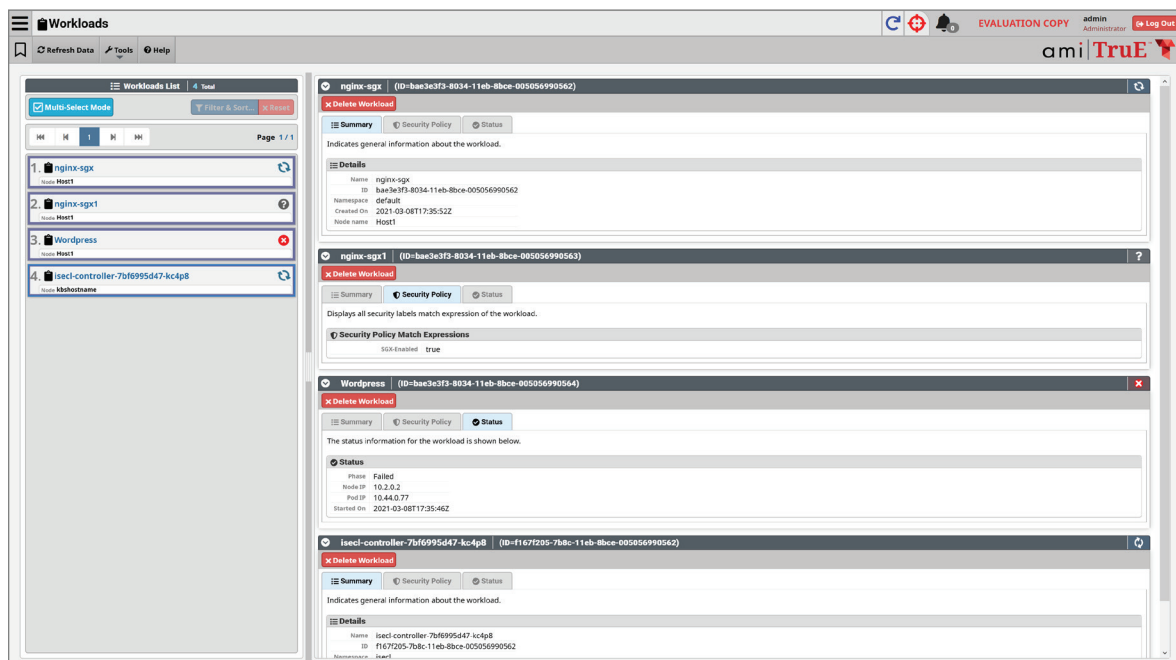


Figure 9. View workloads of a tenant

Summary

SGX attestation is a powerful mechanism that allows data center operators to offer confidential computing as a differentiated security feature and tenants to run their workloads in untrusted environments with the assurance that the confidentiality and integrity of code and data are preserved. AMI TruE and Intel SecL-DC for Data Centers provide the necessary infrastructure components to both data center operators and tenants to take full advantage of SGX attestation.

As a key partner, AMI will continue collaborating with Intel with a shared goal of securing our customers data at rest, in transit and in use.

To learn more, visit www.ami.com.

References

<https://itpeernetwork.intel.com/confidential-computing/>
<https://download.01.org/intel-sgx/latest/>
<https://ami.com/true>
https://download.01.org/intel-sgx/latest/dcap-latest/linux/docs/Intel_SGX_DCAP_ECDSA_Orientation.pdf



¹ Homomorphic encryption also protects data in use. This technology is still maturing and is not covered in this white paper.

² If freshness is desired, the remote relying party can send a nonce to the enclave. The enclave assigns to userData the hash of the enclave-generated data concatenated with the nonce sent by the remote relying party.

Intel disclaims all express and implied warranties, including without limitation, the implied warranties of merchantability, fitness for a particular purpose, and non-infringement, as well as any warranty arising from course of performance, course of dealing, or usage in trade.

Performance varies by use, configuration and other factors. Learn more at www.intel.com/PerformanceIndex. Results have been estimated or simulated. No product or component can be absolutely secure.

The products described may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request.

All product plans and roadmaps are subject to change without notice.

Intel technologies may require enabled hardware, software or service activation. Your costs and results may vary.

© Intel Corporation 2021. Intel, the Intel logo, and other Intel marks are trademarks of Intel Corporation or its subsidiaries. Other names and brands may be claimed as the property of others. Code names are used by Intel to identify products, technologies, or services that are in development and not publicly available. These are not commercial names and not intended to function as trademarks.