Google Cloud

# Protect your keys with Cloud HSM

February 1, 2021

# Table of contents

# Introduction

While Google Cloud encrypts all customer data-at-rest, some customers, especially those who are sensitive to compliance regulations, must maintain control of the keys used to encrypt their data. For these customers, Google Cloud offers Cloud HSM, a service for protecting keys by using a hardware security module.

With Cloud HSM, you can generate encryption keys and perform cryptographic operations in FIPS 140-2 Level 3 certified HSMs. The service is fully managed, so you can protect your most sensitive workloads without worrying about the operational overhead of managing an HSM cluster. The service provides a number of advantages:

- Global availability
- A simple, consistent, and unified API
- Automatic scaling based on your use
- Centralized management and regulatory compliance

Cloud HSM is available in every Google Cloud region around the globe, including multi-regions that span larger geographies. As soon as you enable Cloud HSM, you can create and use HSM-backed keys to protect your data, including data that you store in other Google Cloud services, such as BigQuery, Cloud Storage, and Persistent Disks.

Because Cloud HSM and the HSM hardware are managed by Google, you are spared the complexity and toil of managing and using HSM-backed keys in production, while your data is strictly isolated from other tenants and services in Google Cloud. The Cloud HSM data plane API, which is part of the Cloud KMS API, provides a simplified, intuitive experience.

Cloud HSM supports HSM-backed customer-managed encryption keys (CMEK) wherever CMEK keys are supported across Google Cloud. For example, you can encrypt data in Cloud Storage buckets or Cloud SQL tables by using a Cloud HSM key that you manage.

Cloud HSM is designed to help ensure compliance with regulatory requirements, geographic restrictions, and other business rules. The Security and regulatory compliance section describes how Cloud HSM helps you to verify compliance for existing keys, securely import keys into Cloud HSM, and maintain key provenance.

Finally, the Lifecycle of a Cloud HSM key section shows the sequence of events that takes place behind the scenes when you create a Cloud HSM key, use the key directly, and use it in a CMEK integration.

# Cloud HSM management

Within Cloud HSM, clusters of HSMs are maintained by a team of Google site reliability engineers (SREs) as well as technicians in each Google Cloud data center location. Google handles physical security, logical security, infrastructure, capacity planning, geo-expansion, and data center disaster-recovery planning.

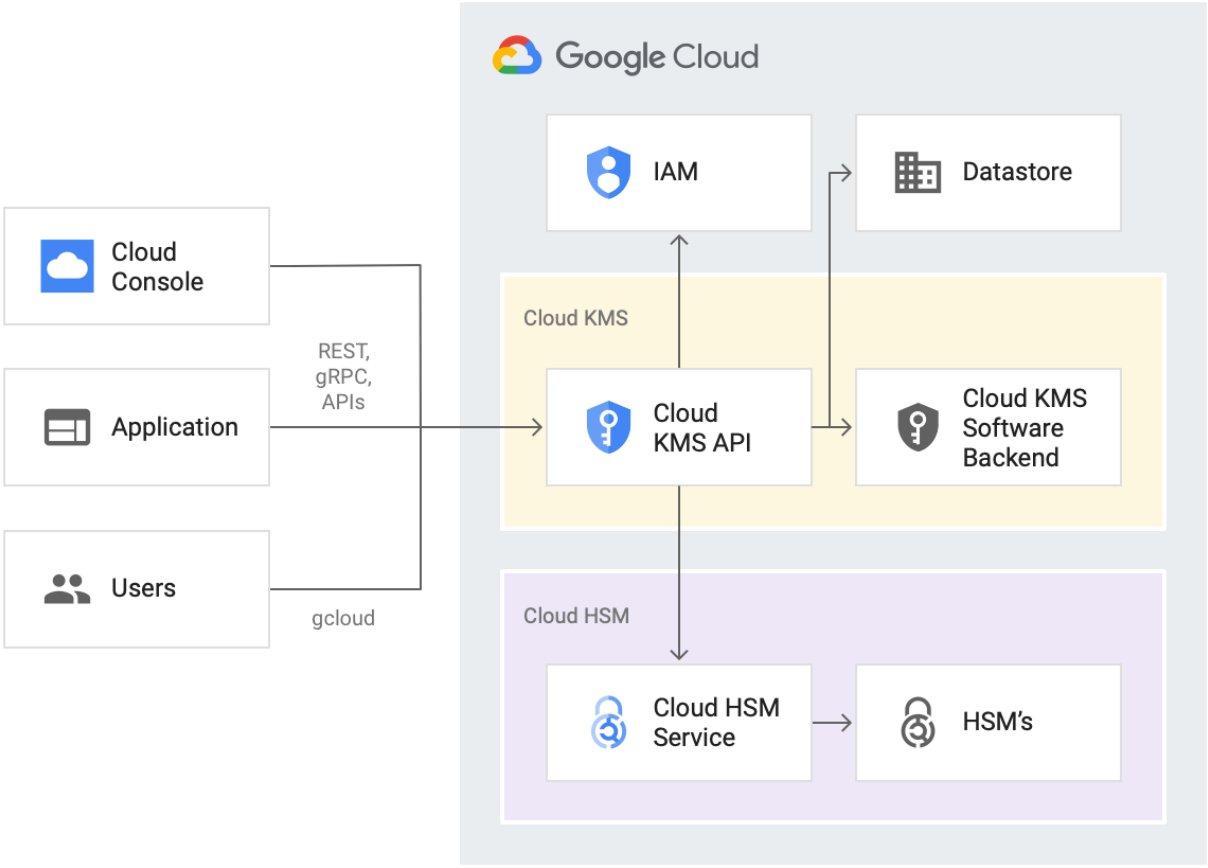
## Abstraction of HSM hardware

Typically, applications communicate directly with HSMs by using both PKCS#11 and a cluster management API. This approach requires you to maintain specialized code for workloads that use or manage HSM-backed keys.

Cloud HSM abstracts away communication with the HSM by proxying requests for HSM-backed keys through the Cloud KMS API. The abstraction reduces the need for HSM-specific code. Cloud HSM inherits Cloud KMS's tight integration with Google Cloud.

Each of these features provides substantial security benefits. The Cloud KMS API significantly reduces the breadth of the HSM interface available, reducing risk in the case of a customer security breach. For example, an attacker would be unable to wipe entire HSMs. Attempts to destroy individual keys are mitigated through a 24-hour safety period.

Cloud KMS's integration with Google Cloud lets you control access to HSM resources using Identity and Access Management (IAM). As a broadly used mechanism, IAM configuration is less likely to suffer from misconfigurations and bugs than a custom HSM solution.



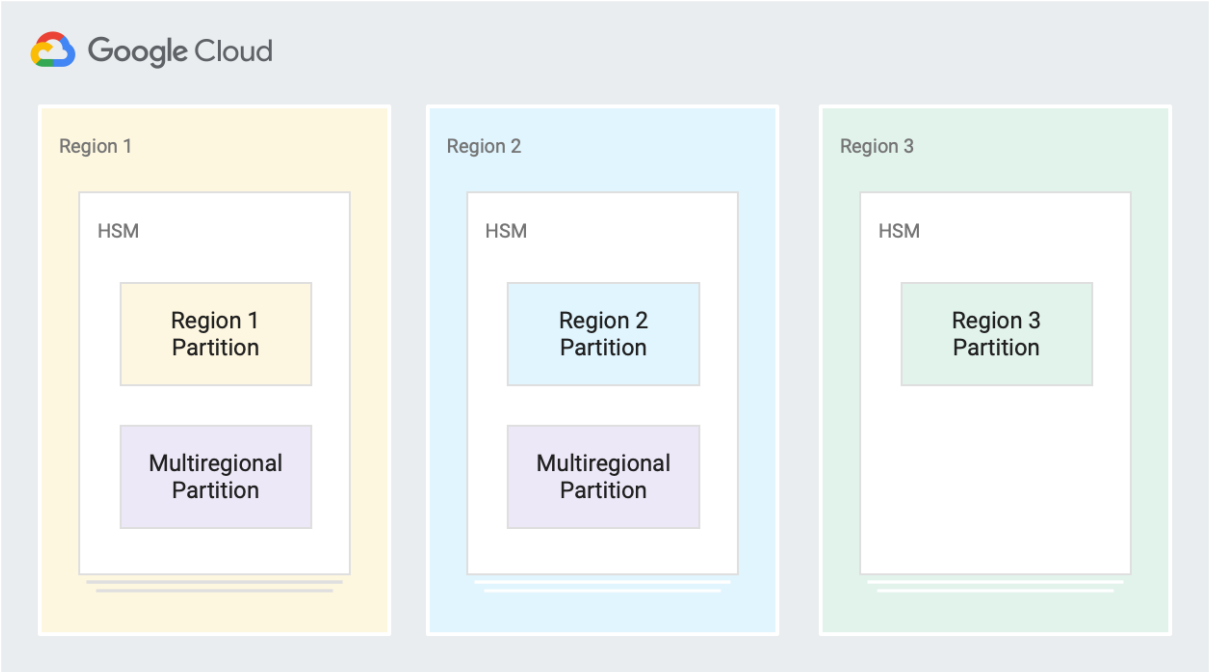

## Strict geographic separation, by design

In Cloud HSM, you can choose to make keys globally available or to enforce strict geographic restrictions on keys that need them.

Often, HSMs are divided into partitions, so that a single physical device can operate as multiple logical devices. You might use partitions to reduce deployment costs in cases where you need to separate HSM administration and keys.

Each Cloud HSM location is associated with a separate root wrapping key. The wrapping key is cloned onto a partition in each HSM in the location, but never leaves the HSM in the location. This allows HSMs in the same region to serve the same set of customer keys, and ensures that HSMs outside the region cannot serve those keys.

Cloud HSM creates multi-regions by cloning a separate wrapping key onto a separate partition on the HSMs in the locations constituting the multi-region. The service leverages the same hardware for multi-regions, but provides the same strong isolation between regions and multi-regions that exists between different regions.




The regionalization scheme hinges on replicating the wrapping keys to only the appropriate partitions. Each configuration change must be approved by multiple members of the Cloud HSM team before it becomes active. Data center technicians can't access an HSM configuration, runtime, or storage in the field.

## Centralized management

In a traditional data center that hosts HSMs, management of the HSMs and their resources is entirely separate from management of other cryptographic resources. Cloud HSM is tightly integrated into Google Cloud, allowing you to seamlessly manage your Cloud HSM resources:

- You manage your HSM-backed resources alongside your other keys in Cloud KMS and externally managed keys in Cloud External Key Manager (Cloud EKM).
- You manage access to HSM-backed resources within IAM.
- Cost reporting for cryptographic operations that use HSM-backed keys are reported in Cloud Billing.

- You can use HSM-backed keys transparently in all Google Cloud services that support encrypting resources using customer-managed encryption keys (CMEK). CMEK integrations require the CMEK key and the data it encrypts to be located in compatible geographic locations. Because of the strict geographic restriction of the Cloud HSM keys, all encryption and decryption of the CMEK data is also geographically restricted.
- Administrative operations on HSM-backed resources are always logged at the API layer in Cloud Audit Logs. You can choose to enable data-access logging as well. For more information, see Cloud KMS audit logging information.
- Google partners directly with the HSM manufacturer to keep the hardware and software on each HSM updated, and to find and fix issues in real time. In the event of a zero-day exploit on the HSM, Google can selectively disable affected code paths on affected HSM clusters until the exploit is fixed.

# Developer and user experience

Because Google is responsible for HSM management, Cloud HSM offers significant benefits to developers and end users.

## HSMs at Google scale

When you rely on hardware that exists on premises or in data centers, the hardware can create a performance bottleneck or become a single point of failure. Cloud HSM is designed to be extremely resilient to unpredictable workloads and hardware failures.

All customer keys are stored wrapped with a regional wrapping key in the KMS database and can only be unwrapped by an HSM in the region as part of a cryptographic operation. This has multiple benefits:

- A key's durability is not tied to a specific HSM or subset of HSMs in a region.
- Each Cloud HSM customer experiences the full scale and availability of the Cloud HSM clusters serving their keys.
- Cloud HSM can handle a much larger set of keys than can be stored on an HSM.
- Adding or replacing an HSM is rapid and secure.

## Unified API design

Cloud HSM and Cloud KMS share a common management and data plane API. The internal details of communicating with an HSM are abstracted from the caller.

Consequently, no code changes are required to update an existing application that uses software keys in Cloud KMS to support HSM-backed keys. Instead, you update the resource name of the key to use.

# Security and regulatory compliance

Cloud HSM helps you enforce regulatory compliance for your workloads in the cloud.

## Cryptographic key attestation

Each time you generate or import a Cloud HSM key, the HSM generates an attestation statement that is signed with a signing key that is associated with the partition. The statement contains information about your key's attributes. The signing key is backed by certificate chains rooted in both Google and the HSM manufacturer. You can download the attestation statement and certificates in order to verify the statement's authenticity and validate properties of the key and the HSM that generated or imported it.

The certificate chain allows you to check the following:

- The HSM hardware and firmware are genuine.
- The HSM partition and HSM are managed by Google.
- The HSM is in the FIPS mode of operation.

The content of the attestation statement allows you to check the following:

- The key is not extractable.
- The key was generated for your CryptoKeyVersion.
- The public key in an asymmetric key pair corresponds to an HSM-backed private key.
- The key material of an imported symmetric key matches the value you wrapped.

## Secure key import directly into HSMs

You can securely import existing keys into Cloud HSM to maintain a backup of your key material outside of Google Cloud, or to simplify migrating certain workloads to Google Cloud. The key-import process does not allow Google any direct access to the unwrapped key material.

Cloud HSM provides you with an attestation statement for the HSM-generated wrapping key to validate that no access occurred.

Because key import potentially creates security and compliance risk by allowing users to bring keys from unknown sources, separate IAM roles allow fine-grained control over who can import keys into a project. Imported keys can be distinguished by the attestation statement that the HSM generates on import. For more information, see Importing a key into Cloud KMS.

## Strict security procedures safeguard HSM hardware

As mandated by FIPS 140-2 level 3, HSM devices have built-in mechanisms to help protect against, and provide evidence of, physical tampering.

In addition to the assurances provided by the HSM hardware itself, the infrastructure for Cloud HSM is managed according to Google Infrastructure Security Design.

Documented, auditable procedures protect the integrity of each HSM during provisioning, deployment, and in production:

- All HSM configurations must be verified by multiple Cloud HSM site reliability engineers (SREs) before the HSM can be deployed into a data center.
- After an HSM is put into service, configuration change can only be initiated and verified by multiple Cloud HSM SREs.
- An HSM can only receive firmware that is signed by the HSM manufacturer.
- HSM hardware is not directly exposed to any network.
- Servers that host HSM hardware are prevented from running unauthorized processes.

## Service and tenant isolation

The Cloud HSM architecture ensures that HSMs are protected from malicious or inadvertent interference from other services or tenants.

An HSM that is part of this architecture accepts requests only from Cloud HSM, and the Cloud HSM service accepts requests only from Cloud KMS. Cloud KMS enforces that callers have appropriate Cloud IAM permissions on the keys that they attempt to use. Unauthorized requests don't reach HSMs.

HSM-backed keys are also subject to quotas for cryptographic operations. These quotas protect your ability to run your workloads by helping to prevent inadvertent or malicious attempts to

overload the service. The default quotas, 3,000 QPM for asymmetric cryptographic operations and 30,000 QPM for symmetric cryptographic operations, are based on observed usage patterns. The quotas are significantly below the service capacity and can be increased upon request.

# Request flows

This section demonstrates how the preceding architectural highlights apply in practice by showing the steps for different types of requests. These flows emphasize the Cloud HSM–specific portions. For more information about steps common to all keys, see the Cloud KMS whitepaper.

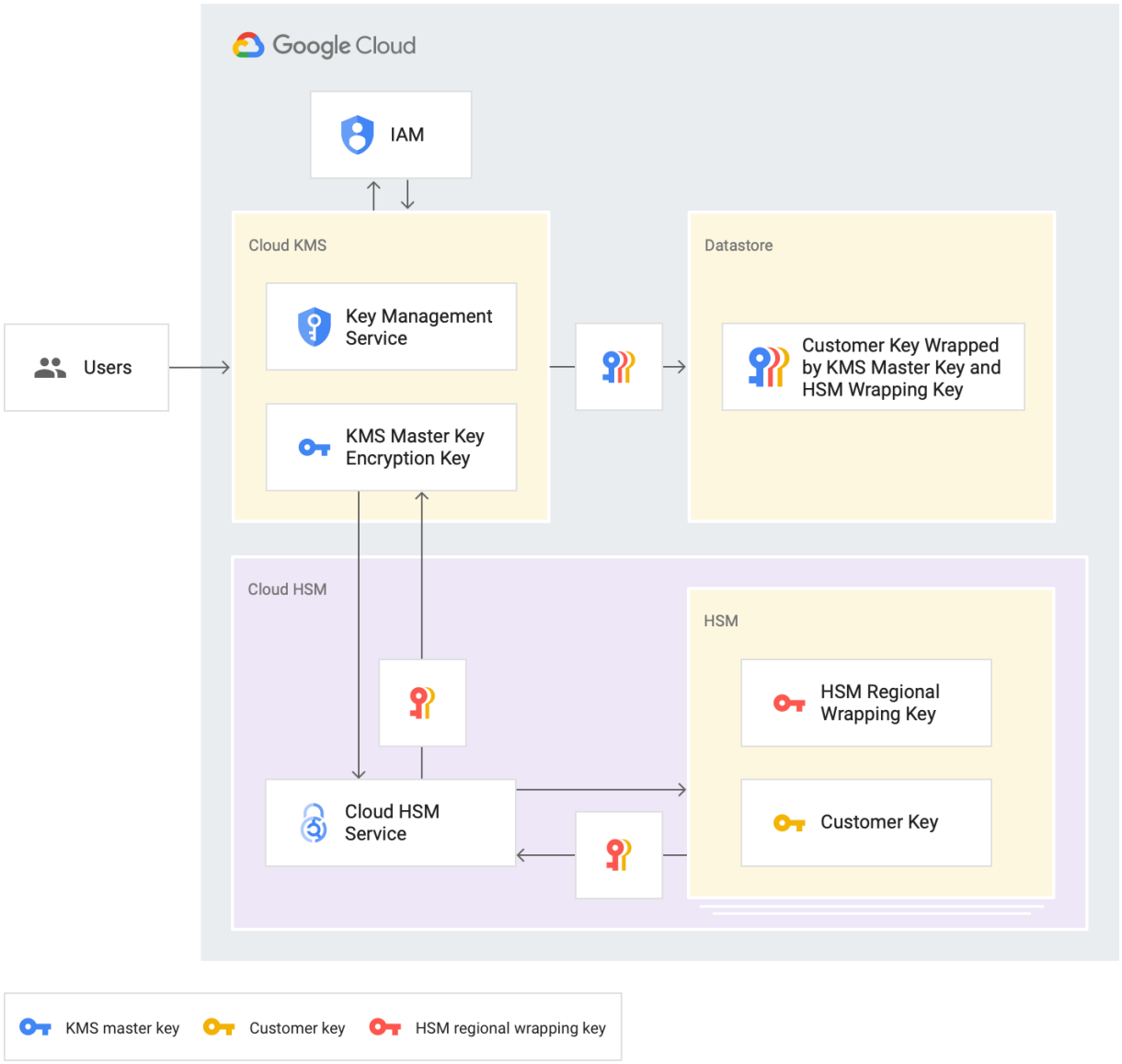
## Creating keys

When you create an HSM-backed key, the Cloud KMS API does not create the key material, but requests that the HSM create it.

An HSM can only create keys in locations it supports. Each partition on an HSM contains a wrapping key that corresponds to a Cloud KMS location. The wrapping key is shared among all partitions that support the Cloud KMS location. The key-creation process looks like this:

1. The Google Front End Service (GFE) routes the key creation request to a Cloud KMS server in the location that corresponds to the request.
2. The Cloud KMS API verifies the caller's identity, the caller's permission to create keys in the project, and that the caller has sufficient write request quota.
3. The Cloud KMS API forwards the request to Cloud HSM.
4. Cloud HSM directly interfaces with the HSM. The HSM does the following:
    a. Creates the key and wraps it with the location-specific wrapping key.
    b. Creates the attestation statement for the key and signs it with the partition signing key.
5. After Cloud HSM returns the wrapped key and attestation to Cloud KMS, the Cloud KMS API wraps the HSM-wrapped key according to the Cloud KMS key hierarchy, then writes it to the project.

This design ensures that the key can't be unwrapped or used outside of an HSM, can't be extracted from the HSM, and exists in its unwrapped state only within locations you intend. The following diagram shows the differences when creating Cloud HSM keys and software keys in Cloud KMS.



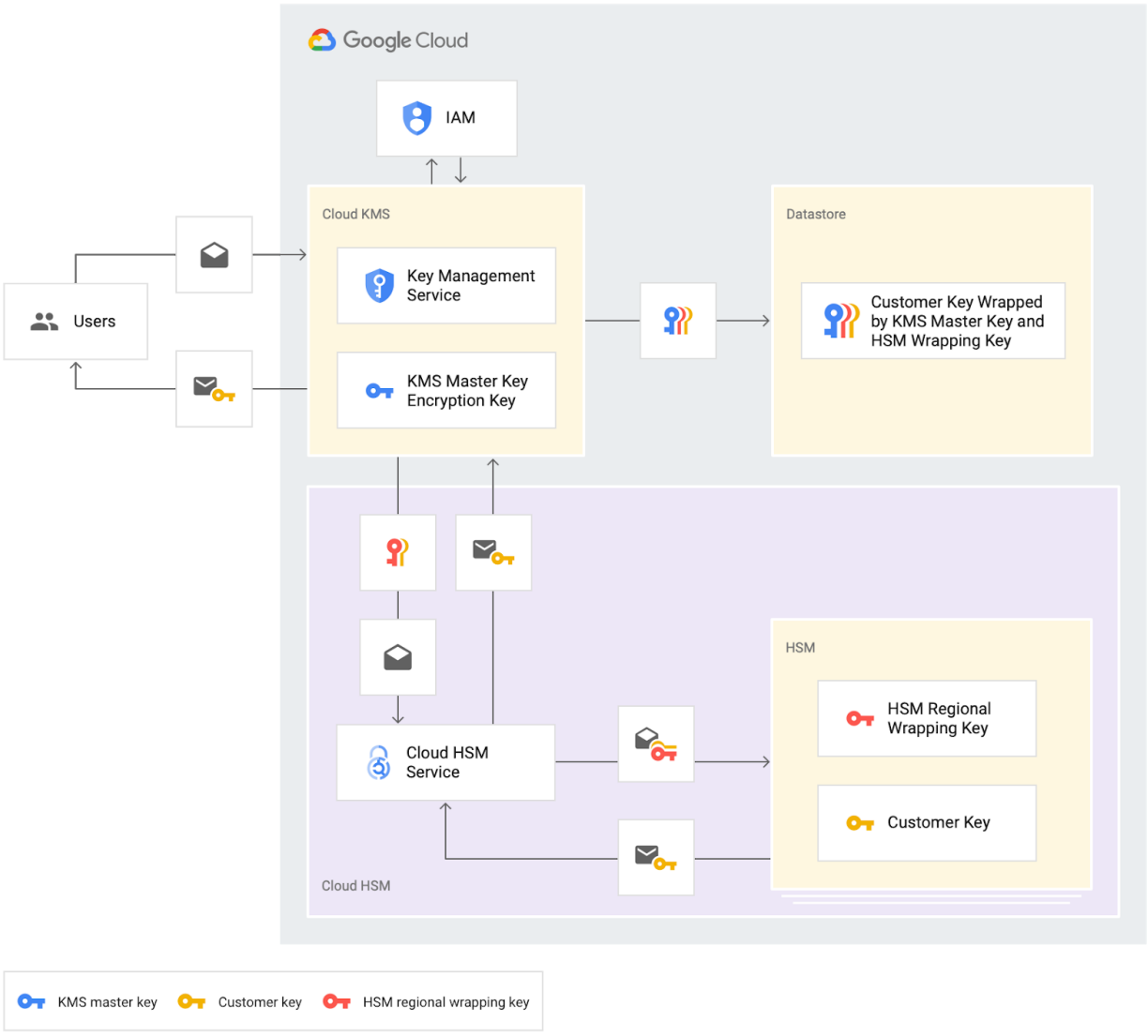

## Cryptographic operations

When you perform a cryptographic operation in Cloud KMS, you don't need to know whether you are using an HSM-backed or software key. When the Cloud KMS API detects that an operation involves an HSM-backed key, it forwards the request to an HSM in the same location.

1. The GFE routes the request to a Cloud KMS server in the appropriate location. The Cloud KMS API verifies the caller's identity, the caller's permission to access the key and perform the operation, and the project's quota for cryptographic operations.
2. The Cloud KMS API retrieves the wrapped key from the datastore and decrypts one level of encryption using the Cloud KMS master key. The key is still wrapped with the HSM wrapping key for the KMS location.
3. The Cloud KMS API detects that the protection level is HSM and sends the partially unwrapped key, along with the inputs to the cryptographic operation, to Cloud HSM.
4. Cloud HSM directly interfaces with the HSM. The HSM does the following:
    a. Checks that the wrapped key and its attributes have not been modified.
    b. Unwraps the key and loads it into its storage.
    c. Performs the cryptographic operation and returns the result.
5. The Cloud KMS API passes the result back to the caller.

Cryptographic operations using HSM-backed keys are performed entirely within an HSM in the configured location, and only the result is visible to the caller.

This diagram shows the difference between creating Cloud HSM keys and software keys in Cloud KMS.




## CMEK integrations

CMEK and Cloud HSM together allow you to protect your data in select Google Cloud services with HSM keys. Configuring a CMEK-enabled service to use Cloud HSM keys is as simple as choosing a key with an HSM protection level when following the service-specific instructions. When a caller reads or writes data to a CMEK-enabled service, the caller doesn't need direct permission to use the key, nor does the caller need to know whether the key is stored in an HSM.

The flow for a CMEK operation is similar to that for a normal cryptographic operation with the following exceptions:

- The request from the CMEK-enabled service is initiated within Google's network, and does not need to traverse the GFE.
- The Cloud KMS API verifies that the service account for the CMEK-enabled service has proper permissions to use the key. The Cloud KMS API does not validate permissions on the end-user of the CMEK-enabled service.

Cloud HSM is Google Cloud's hardware key management service. It offers a number of distinct advantages to users looking to protect their data-at-rest with HSM keys. The service was designed with the principles of locked-down API access to the HSMs, effortless scale, and tight regionalization of the keys.

Cloud HSM offers CMEK support for the most important services and presence of Cloud HSM in every Google Cloud region (including multi-regions and global). The service is designed to make it easy for you to protect your sensitive data, wherever it may be, with a key that's protected by FIPS 140-2 Level 3 devices.

To comment on this paper, contact us at cloudhsm-feedback@google.com.

# Further reading

To learn more, explore the following resources:

- Cloud HSM documentation
- Cloud KMS documentation

Whitepapers:

- Cloud KMS deep dive
- Google security
- Google infrastructure security design overview
- Trusting your data with Google Cloud
- Data encryption at rest
- Data encryption in transit
- Data deletion on Google Cloud

Other documentation:

- Binary Authorization for Borg (BAB) (code provenance)
- Access Transparency
- Government requests for customer data: controlling access to your data in Google Cloud
- Data residency, operational transparency, and privacy for European customers on Google Cloud
- Identity and Access Management
- Cloud Audit Logs
- Cloud Asset Inventory