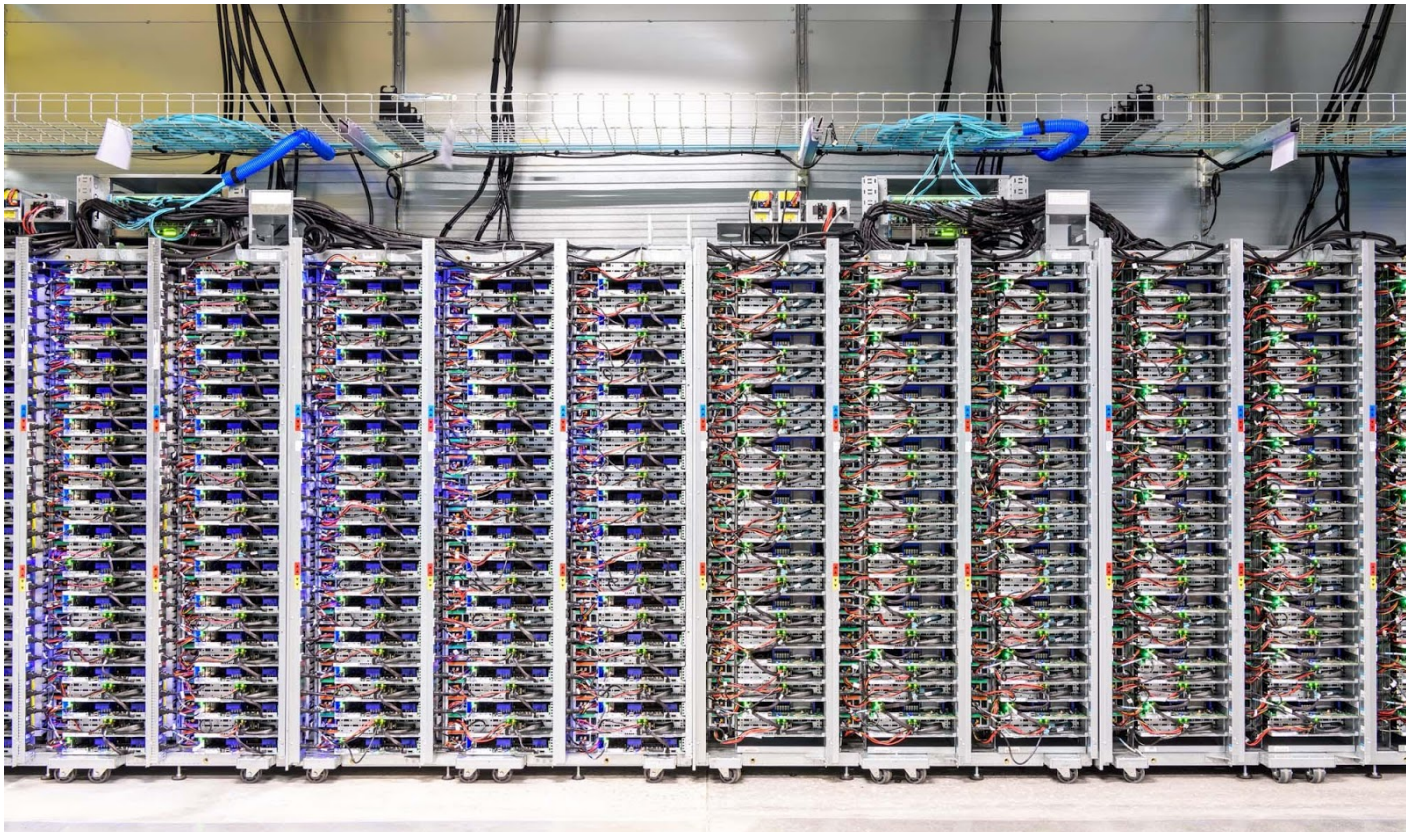
Google Cloud Whitepaper
March 2021

Protecting data with Google Cloud External Key Manager



Authors:

Andrew Lance, Founder & Principal, Sidechain
Dr. Anton Chuvakin, Head of Solutions Strategy, Google

Google Cloud

Table of contents

Disclaimer

The content contained herein is correct as of March 2021, and represents the status quo as of the time it was written. Google Cloud's security policies and systems may change going forward, as we continually improve protection for our customers.

Establishing greater trust in the cloud

Cloud adoption has far surpassed being a mere technology solution -- Cloud platforms now function as vast utility infrastructure, adopted by organizations of all sizes. Organizations are leveraging these cloud services to accelerate their ability to bring applications to market without the overhead of maintaining their own infrastructure. Cloud platforms like Google Cloud Platform enable access to world-wide scalability, resilience, and an incredible array of integrated services and security capabilities.

Despite these benefits, one of the biggest barriers to cloud adoption is data security. While organizations are continuing to entrust their data to cloud providers, security concerns persist, particularly for sensitive or regulated data. Ultimately, relying on a cloud service provider (CSP) to store and process data requires trust and a willingness to give up some degree of control that was maintained when organizations managed their own infrastructure within their own data centers. And while the hypergrowth and adoption of cloud platforms has been astonishing -- by 2022, public cloud services spend is expected to rise to $362.3 billion, a 66% increase from 2019 -- there are a variety of security concerns that continue to block wholesale cloud adoption, particularly when sensitive data is involved.

In addition, geopolitical realities mean that governments are increasingly asserting the need for sovereignty over data and infrastructure, driving cloud providers to deploy regional solutions and customers to architect with regional dependencies in mind. For example, since US cloud providers, which most EU organizations use, are governed under regulations that enable US authorities to access data stored within their infrastructure (such as the CLOUD Act), there remains a concern putting highly sensitive data in such cloud platforms.

An increasing number of countries are also establishing data privacy and protection laws that often require data created in-country to remain there. The General Data Protection Regulation (GDPR), which entered into force in 2018, imposes rules for the processing of personal data and how that data is transferred both within the EU and to non-EU countries. In some cases, individuals and organizations may risk losing protections of local data privacy regulations if their data is transferred out of the country. Cloud providers that operate globally need to give customers a way of managing the regionality of their data as it's created, used, and stored.

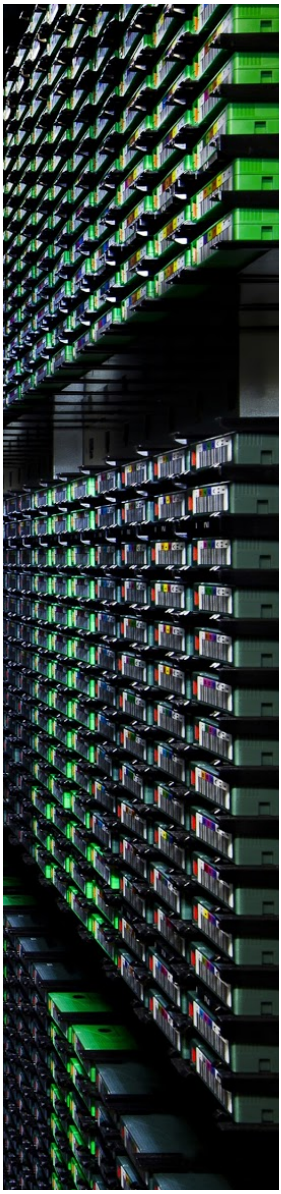
Organizations also worry about data safeguards within cloud platforms. After all, if an organization considers "cloud" to be just running your workloads on someone else's computers, there is a legitimate concern about who has access to those computers, particularly when cloud computing platforms are inherently multi-tenant. Google runs a world class security program, that is likely on par or better than that of other leading organizations. Even this, however, isn't complete assurance for highly regulated companies that need to attest to the very highest security standards that they adopted. Even when a cloud platform like Google Cloud has robust security, it does not automatically mean that it will be trusted by everybody.

To address these concerns, Google has developed a host of services that deliver high-assurance security capabilities that enable customers to better safeguard their data and reduce the amount of trust they need to place towards the Google Cloud. Confidential Computing, for example, enables customers to encrypt data in-use while it's being processed. By using Confidential VM's and Confidential Google

Kubernetes Engine (GKE) nodes, data can remain protected and encrypted at all times while in use, all without making any code changes or compromising on performance.

Cloud External Key Manager (EKM) is another ground-breaking capability that enables customers to encrypt data in a variety of services including BigQuery, Compute Engine, Cloud SQL, and Google Kubernetes Engine with encryption keys that are stored and managed in a third-party key management system deployed outside Google's infrastructure. EKM maintains separation between encrypted data in Google Cloud and encryption keys stored outside the cloud. Combined with Key Access Justifications, which provides a justification for every encryption key request, customers can now be the ultimate arbiter of access to data within the Google Cloud.

Key security - giving customers the choice



Google Cloud has long offered several different levels of customer control for data encryption and key management. Google Cloud was the first cloud platform to offer data-at-rest encryption by default, ensuring that all data stored within the cloud is encrypted by Google-managed keys. Without any key management overhead whatsoever, customers can rely on the fact that all of their data is encrypted at rest transparently and securely.

For those customers that need more control over keys, Google Cloud offers Cloud Key Management Service (KMS) and Customer Managed Encryption Keys (CMEK) . These capabilities not only provide the mechanisms to create keys of various types and strengths, but also the region in which the keys are created and stored, the ability to rotate keys, and of course, disabling and destroying keys as well.

Google Cloud also offers Cloud HSM, a cloud-hosted Hardware Security Module (HSM) service that hosts encryption keys and performs cryptographic operations in a cluster of FIPS 140-2 Level 3 validated HSM's. Google manages this HSM cluster, so normal operational tasks such as redundancy, scaling, and patching are automatically provided. And because Cloud HSM uses Cloud KMS as its front end, HSM-backed keys can be used as a CMEK key anywhere a KMS software key can be used.

Cloud External Key Manager (EKM), however, is a dramatic step towards giving customers ultimate control over their keys and encrypted data-at-rest within Google Cloud. In yet another first, Google EKM enables customers to use keys managed in a supported external key management system to protect data within Google Cloud.

Like Cloud HSM, Cloud EKM leverages the Cloud KMS as the frontend and, as a result, customers can use the same UI and API to create and use an EKM key. In this way, Google Cloud has offered a unified way of selecting a number of key management options all within a single interface. Whether you choose to manage keys in Cloud KMS directly, leverage hardware keys through Cloud HSM, or choose Cloud EKM for external key management.

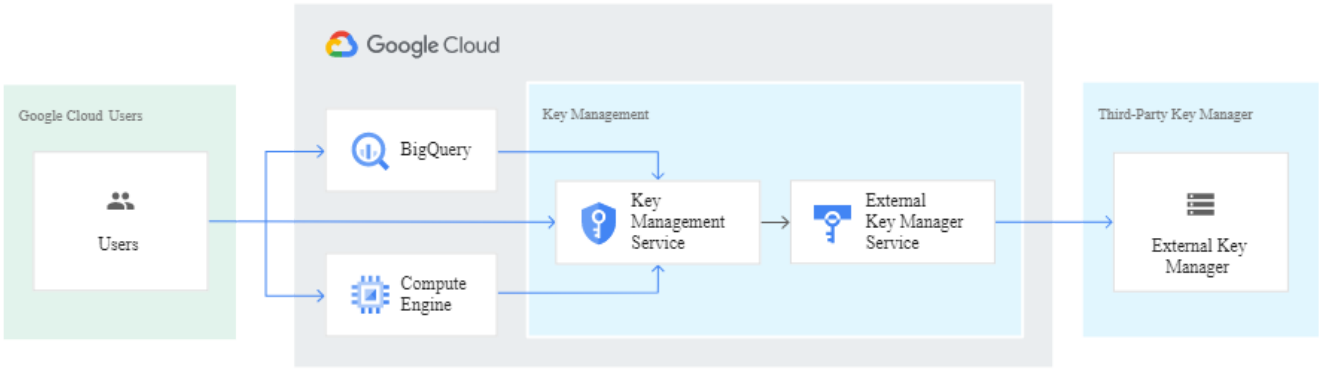
Cloud EKM offers customers several unique advantages and benefits to key management that are unavailable with other cloud platforms.

What is Cloud EKM and how does it work?

Cloud EKM enhances data protection in Google Cloud by empowering customers to maintain control of when cloud services can access encrypted data-at-rest. Without the customer explicitly authorizing access to the data, Google Cloud services are blocked from decrypting customer data. How does the customer *allow* cloud services to access encrypted data? There are 3 steps for creating an EKM key:

1. First, the customer creates or uses an existing key in a supported external key management partner system. This key has a unique URI. Because this key is created and managed by the customer, Google is not involved in any way in the management of this key.
2. Next, the customer grants Google Cloud access to use that key stored in the external key management system. This is done by both configuring your Google Cloud project with the URI information of the key, as well as through the external key management system to grant access to that target URI.
3. In Google Cloud, the customer creates a Cloud EKM key, using the URI for the externally-managed key.

The following diagram shows an overview of how Cloud EKM works:



In summary, all customer data configured in this way is protected by a Cloud EKM key. But in order for Google to decrypt customer data, it must have access to the externally-managed key.

When Google makes this key request, and if configured to do so, it also must provide a Key Access Justification -- metadata that describes why the request is being made -- giving customers the ability to also make a policy decision about whether to grant access to the external root key. If, for any reason, the customer determines that the key request is not justified, they can simply block the key request, thus preventing Google Cloud from decrypting customer data.

These powerful controls, such as requiring Google Cloud to decrypt data only by accessing a customer-controlled root key and requiring Google Cloud to provide a key access justification (if configured to do so) from which the customer can determine whether to fulfill the request both form the bedrock for giving customers control over how Google accesses their sensitive data-at-rest in the cloud.

Primary benefits of Cloud EKM

While there are many reasons a customer may want to take advantage of this powerful separation of duties to protect their data in the cloud, there are three primary drivers customers cite when adopting Cloud EKM.

Key provenance

Key provenance is evidence that details the origin, changes to, and supporting confidence or validity of encryption keys such that customers can reason about the origin, location, backup history, and other characteristics of the key. Example metadata include when the key was created, by which cryptographic engine, authorized by whom, how it was generated, and under what circumstances. In some cases, elaborate *key ceremonies* take place that capture key creation processes to precise details, and secure the output of those keys in various ways.

Key provenance is not a one-time event at key creation time. Key provenance must also provide evidence about how keys are stored, accessed, used, and destroyed along its entire lifecycle. Provenance provides a platform for evaluating **trust**, as well as other qualitative metrics that enable customers to use the cloud with confidence.

Customers often have provenance requirements for data encryption keys for a variety of reasons. These keys usually protect sensitive and/or regulated data. The effectiveness of the protection is a function of the efficacy of the security control, in this case, an encryption key. If an organization is accountable to a compliance audit, governing body, federal regulation, or security policy that requires adequate protections are in place to safeguard such data, using provenance evidence is one powerful way of attesting to the effectiveness of the encryption.

This creates a conundrum for some organizations moving to the cloud with strict requirements in this area. After all, encrypting data with a key that was produced in some “unknown” way by a third-party (i.e. cloud provider) may not give the organization the provenance evidence they require. Likewise, while leveraging a cloud platform for key management is certainly operationally efficient, and even satisfying rigorous security inquiry, such as how the keys are stored, a comprehensive access model about who has access to the key platforms, and clear evidence of when a key is invoked can be solved. However, there are clients who are not satisfied by anything of this sort - they just have to have the encryption keys in their physical control and full possession, as well as under sole administrative control.

One thing customers value in cloud platforms is the redundancy and “always available” nature of data, which means in some cases, vast replication of data to achieve this high availability, and most cloud-based key management services are no different. A cloud platform may provide assurance that once a *key deletion* operation is invoked, that key is absolutely destroyed across all instances and copies of that key within the cloud infrastructure, and that all such key operations were logged and audited. However, what if the client security team does not trust the assurance? Or, there are other requirements that can only be satisfied if the client does it on their own. In these cases, such concerns are mitigated by simply removing access via EKM.

Key centralization

Centralized key management has the following advantages:

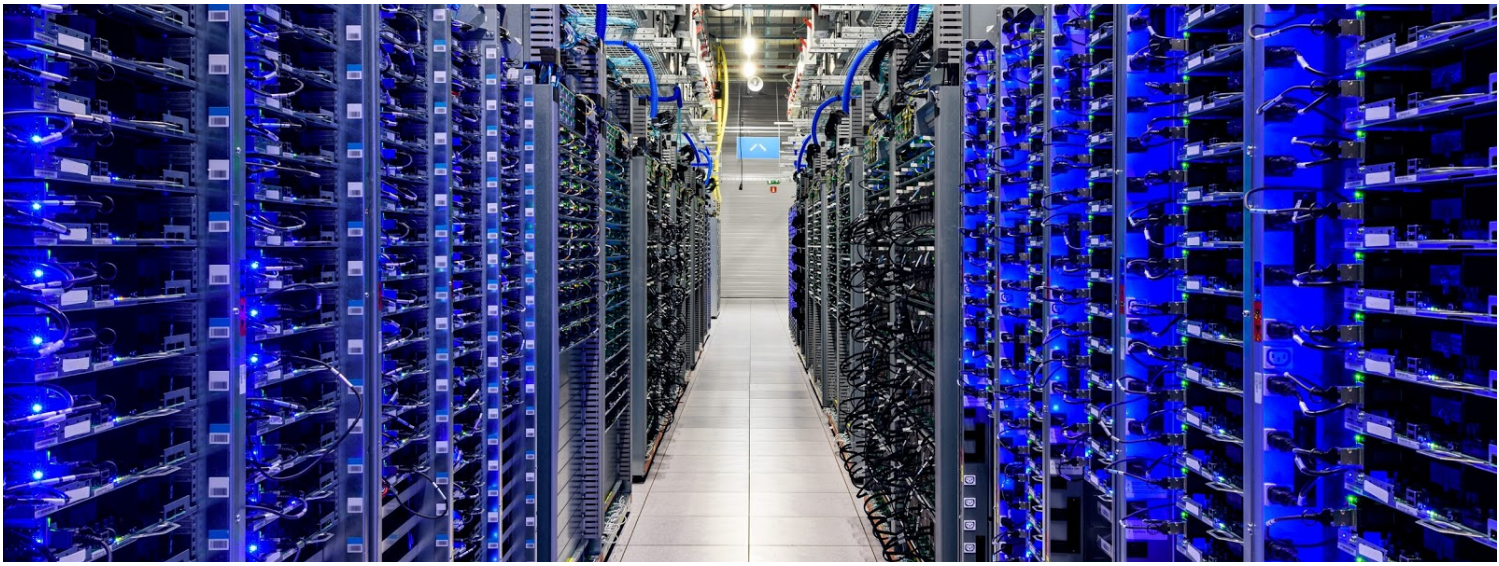
- Fewer key management systems means less distributed risk
- One KMS to manage, easier to operationalize and keep staff trained on
- Enables a multi-cloud strategy, using keys within different platforms from a single source
- Centralized governance, provenance, and audit of keys

For organizations with sophisticated key management infrastructure, centralizing key management functions is a critical architectural and procedural goal that greatly improves the ongoing operations and management of those systems.

Less distribution of keys means less distributed risk of a compromised key. By centralizing key management infrastructure, it reduces the threat landscape where keys are a target, enabling organizations to apply greater security to a smaller subset of infrastructure. Centralization enables teams to focus the governance and audit function to a smaller set of systems, making audits and security review much more efficient. In fact, businesses go to great lengths to reduce the number of systems that fall under audit requirements where keys are concerned, for example in PCI, by tokenizing regulated data elements. Teams can also focus on developing a smaller set of operational skills for a centralized key management system than having to manage several KMS's.

Centralized key management also enables flexibility in architecture, such as achieving multi-cloud strategies. When primary key management functions are maintained on a centralized system - key creation, key rotation, key deletion, for example - those keys can enable cryptographic functions in any number of cloud platforms, without increasing much overhead. It also enables hybrid architectures that leverage existing key management systems while encouraging data to be used in powerful cloud platforms like GCP.

Centralization delivers the best of both worlds, enabling on-premises control of keys through tried-and-true key management infrastructure while still leveraging the cloud for data processing and scalability.



Key control

The geopolitical concerns around key management and data protection are very real, particularly as the European Union, Latin America, and African countries grapple with how to maintain sovereignty over in-country data when that data is managed by cloud platforms concentrated in the hands of American or Chinese cloud providers. After all, using a cloud provider to manage business infrastructure, data, keys, and policies is putting a lot of eggs in one basket. Cloud EKM, and a host of other GCP features and services, are giving control back to customers, and helping earn more customer trust because of it.

As a result, Cloud EKM gives customers the ability to:

- **Protect that key** - it's encrypted by a customer managed key that is in customer possession, likely in their physical location and under their administrative control
- **Authorize the use of that key** - through a request to invoke the customer managed key
- **Revoke access to that key** - by shutting off access to the customer managed key
- **Validate the use of that key** - through Key Access Justification metadata
- **Maintain complete provenance** of a root key that protects the entire data access model

By extending this level of control to customers, Google Cloud is asking for *less trust*, because customers no longer are required to give up control over their most sensitive data-at-rest. If, for any reason, be it geopolitical, regulatory, or otherwise, a customer of Google Cloud wishes to revoke access to their data, they may do so, and do it easily. The process doesn't require arduous workflows, moving massive amounts of data around, or initiating the deletion of vast customer data in the cloud. It's achieved by simply removing cloud access to the key: disabling the key in the customer key manager, shutting off or blocking network access, making an access policy change, or any number of other options.

Google Cloud is working to earn trust not by hand-waving or demanding complete control over the processing of customers' most sensitive data. It is doing it by allowing customers to retain control, empowering customers, and limiting its exposure to sensitive customer data while still providing stellar value through cloud services.

Core use cases

While there are many examples of how customers use Cloud EKM to protect their data in Google Cloud, the following three examples are common use cases that customers cite for leveraging this capability. Additional discussion of these can be found in this blog.

Protecting highly sensitive data

Not all data is created equal. As cited before, data is usually classified into categories of sensitivity. Organizations often classify their applications and workloads in the same way. Moving non-sensitive workloads to the cloud usually doesn't demand the highest level of security controls, and doesn't normally have the highest compliance requirements. In fact, customers of the cloud have often started by moving these kinds of non-critical workloads for these very reasons.

As higher sensitivity workloads are moved to the cloud, greater security controls are required and audit requirements increase. Google Cloud offers a high level of security controls and configuration for data and these may be suitable for most data managed by organizations moving to the cloud.

Some organizations, however, will run into the "last 20%" -- a way of capturing the highest level of sensitive data that organizations are, understandably, hesitant to move to the cloud. Even if the business and economic drivers are there, the security and compliance risks are holding these workloads back.

Cloud EKM helps unlock such "reticent" workloads by providing customers with yet a greater level of control than previously offered, for all the reasons cited above. It also provides protections against:

- Accidental misconfiguration of cloud security controls: if the cloud provider misconfigured key access settings, it could result in key disclosures, however with Cloud EKM, the provider does not have custody of the key
- A rogue employee of the provider can never access encryption keys because the provider does not have the keys - they are stored at a client site.
- If an entity requests that a provider surrender the keys to some particular customer data, this is impossible because the keys are not in the provider's possession

Retain control to address geopolitical and regional concerns

Data protection laws are estimated to increase from covering 10% of the world's population to 65% by 2023, according to Gartner. These regulations involve how personal data is handled, shared, transferred, and protected, and will (if they follow the lead of current regulations) levy significant fines if companies are found to be noncompliant.

With emerging regulatory guidelines that regulate how data is transferred between country borders and where regulated data is stored, it poses a challenge for international customers that the large cloud providers are concentrated in the U.S. and China.

One strong step that EU-based customers can take is to maximize the control they have over data put into the cloud. Cloud EKM offers a maximum level of protection, control, and ability to "shut off" access in the face of a changing regulatory landscape (via Key Access Justifications).

Support hybrid and multi-cloud architectures

Large enterprises, particularly within highly regulated industries, maintain sophisticated key management systems that involve more than just a large array of HSMs and other systems to secure encryption keys. They have extensive policies that define the criteria for handling these systems and other cryptographic material, as well as processes that define staff roles, activities, procedures, and operational support for these systems.

For these organizations, using keys stored with a cloud provider like Google represent *at best* a significant risk both operationally and to audit and compliance and incredible overhead to risk management as a result. Worst case, it's an impossibility, which blocks the utilization of valuable cloud services and resources.

Cloud EKM enables these customers to leverage their current key management infrastructure (provided one of the supported partners is used) to maintain key provenance while still putting protected workloads in the cloud. Google Cloud will never store this key, and isn't responsible for creation or deletion since those tasks are maintained by the customer.

Service integrations and technical considerations

Cloud EKM is available for several services running in Google Cloud. At the time of this writing, the following services support Cloud EKM keys:

- **Compute Engine/Persistent Disk**
  - By default, Compute Engine encrypts customer data at rest. And as part of this default encryption, Compute Engine handles and manages this encryption for you. You can, however, control and manage this encryption by using Cloud EKM keys. In this case, the Persistent Disks used with your instances, along with images and snapshots, are encrypted with data encryption keys protected by your Cloud EKM key.
- **BigQuery**
  - Data stored in BigQuery can be encrypted using Cloud EKM keys. Setting this up is similar to configuring BigQuery with other customer managed encryption keys, except in this case, use a Cloud EKM Key. Even data in BigQuery cache will need access to the customer-managed key before executing, enabling data to be instantly revoked from BigQuery by the customer.
- **Google Kubernetes Engine: Data on VM disks or Application-layer Secrets**
  - In GKE, Cloud EKM keys can be used to protect data of two types of storage disks: node boot disks and attached disks. In addition, Cloud EKM keys can be used to protect Kubernetes Secrets at the application layer.
- **Cloud SQL**
  - Cloud EKM keys can be used to encrypt Cloud SQL instances. Backups of these instances are also encrypted with the same key.

It's important to note that Google does not store your keys on its servers and cannot access your protected disks unless you provide the key to Cloud EKM. This also means that if you lose this key, or access is lost, there is absolutely no way for Google to recover the key or to recover any data encrypted with the lost key.

Another important caveat to using Cloud EKM is that since Google Cloud is making a remote connection to the customer-managed key infrastructure, it's usually wise to use Cloud EKM key rings in a region close to the external key management system. This minimizes the latency of this remote call.

Integration solution providers

Google Cloud supports a range of integration solution providers to enable Cloud EKM. The following external key management partner systems are supported:

- Fortanix
- Ionic
- Thales
- Equinix SmartKey
- Unbound Tech

Fortanix

Implemented as a feature within the Fortanix Self-Defending KMS (SDKMS), customers can utilize the Fortanix solution to hold their own keys enabled by Cloud EKM. The solution is available on-premises with a FIPS 140-2 Level 3 validated Fortanix Runtime Encryption Appliance, or as software that can be deployed on-premises.

Ionic

Ionic Machina enables customers to create, use, and store their own encryption keys integrated with Cloud EKM. It enables real-time policy enforcement, a single unified view on how data is accessed within GCP, and attribute-based access controls using Google Key Access Justifications.

Thales

The CipherTrust Key Broker by Thales integrates with Cloud EKM to enable customers to implement separation of duties, controlling keys separate from their sensitive data in the cloud. The industry-leading Thales Luna Cloud HSM acts as the trust anchor for the CipherTrust Key Broker, and is a FIPS 140-2 Level 3 certified root-of-trust.

Equinix SmartKey

SmartKey, powered by Fortanix, provides SaaS-based secure key management for Cloud EKM and is available globally in multiple regions including United States (AMER), United Kingdom (UK), European Union(EU), Asia Pacific (APAC) and Australia (AU). Keys are automatically replicated within-region to create high availability access.

Unbound

Cloud EKM keys can be used outside of Google Cloud Platform using Unbound Key Control (UKC), the first and only software-based solution FIPS 140-2 Level 2-certified by NIST. UKC leverages secure multiparty computation (MPC) to protect cryptographic keys by ensuring they never exist in their complete forms throughout their lifecycle.

Unlock new cloud workloads with Cloud EKM

Enterprises are finding unparalleled value by shifting infrastructure and workloads to the cloud. As Google introduces features and services that continue to drive these efficiencies, it's imperative that security capabilities accompany them to ensure that workloads and data in the cloud are safer than ever. In a monumental shift, Google has pioneered the model of creating *more trust* by actually trusting the cloud provider less.

By enabling customers to control the managed key that cryptographically limits how Google can access customer sensitive data, Google Cloud now has established the opportunity to make permission-based access standard across an array of services. Whether it's to prove provenance of the root-of-trust for encrypted data in the cloud, to further wrap security around the most sensitive workloads, or even to address geopolitical concerns between countries across the globe, Google Cloud EKM keys offer an unmatched level of security control and trust for customers of GCP.