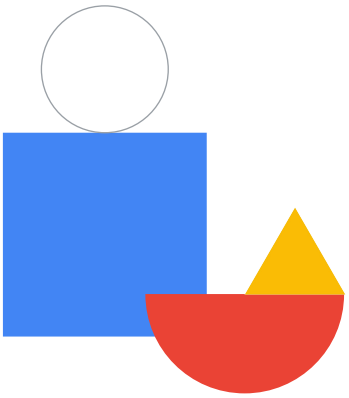
Google Cloud

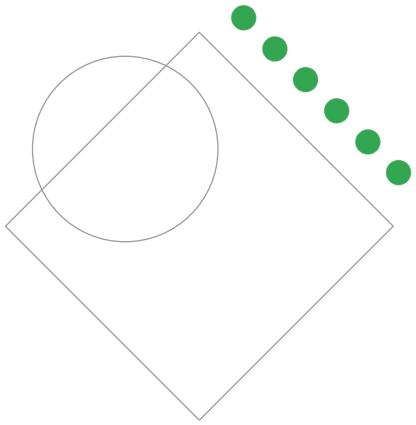


Preparing for Your Professional Cloud Security Engineer Journey

Module 3: Ensuring Data Protection

Welcome to Module 3: Ensuring Data Protection.

Review and study planning

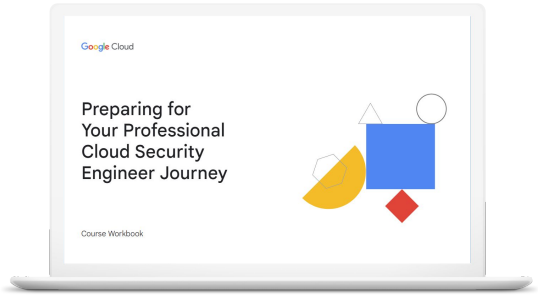


Google Cloud

Now let’s review how to use these diagnostic questions to help you identify what to include in your study plan.

Your study plan:

Ensuring data protection



3.1	Protecting sensitive data and preventing data loss
3.2	Managing encryption at rest, in transit, and in use
3.3	Planning for security and privacy in AI

Google Cloud

We'll approach this review by looking at the key areas of this exam section and the questions you just answered about each one. We'll talk about where you can find out more about each area in the learning path for this certification and/or where to find the information in Google Cloud documentation. As we go through each one, take notes on the specific courses (and modules!), skill badges, and documentation pages you'll want to emphasize in your study plan.

3.1 Protecting sensitive data and preventing data loss

Considerations include:

- Inspecting and redacting personally identifiable information (PII)
- Ensuring continuous discovery of sensitive data (structured and unstructured)
- Configuring pseudonymization
- Configuring format-preserving encryption
- Restricting access to BigQuery, Cloud Storage, and Cloud SQL datastores
- Securing secrets with Secrets Manager
- Protecting and managing compute instance metadata

Google Cloud

As a Professional Cloud Security Engineer, you play a critical role in protecting sensitive data. This includes being able to configure and run data-loss prevention (DLP) software such as Sensitive Data Protection. It also includes protecting sensitive data with encryption, access control, service perimeters, and secure techniques for data manipulation with such Google Cloud tools as VPC Service Controls and Secrets Manager.

Question 1 tested your knowledge of configuring DLP to automatically inspect and redact personally identifiable information (PII). Question 2 tested your knowledge of using Sensitive Data Protection features. Question 3 asked you to secure access to BigQuery and question 4 explored your understanding of Secrets Manager.

3.1 Diagnostic Question 01 Discussion

Cymbal Bank has hired a data analyst team to analyze scanned copies of loan applications. Because this is an external team, Cymbal Bank does not want to share the name, gender, phone number, or credit card numbers listed in the scanned copies. You have been tasked with hiding this PII information while minimizing latency.

What should you do?

A. Use the Cloud Data Loss Prevention (DLP) API to make redact image requests. Provide your project ID, built-in infoTypes, and the scanned copies when you make the requests.

B. Use the Cloud Vision API to perform optical code recognition (OCR) from scanned images. Redact the text using the Cloud Natural Language API with regular expressions.

C. Use the Cloud Vision API to perform optical code recognition (OCR) from scanned images. Redact the text using the Cloud Data Loss Prevention (DLP) API with regular expressions.

D. Use the Cloud Vision API to perform text extraction from scanned images. Redact the text using the Cloud Natural Language API with regular expressions.

Google Cloud

Feedback:
A. Correct! The DLP API can be directly used for image redaction. Built-in infoTypes already include name, gender, phone number, and credit card numbers.

B. Incorrect. The Cloud Vision API's OCR can be used to extract text from images but OCR does not redact the text. The Cloud Natural Language API also cannot help with text redaction. Use the DLP API for this.

C. Incorrect. The Cloud Vision API's OCR can be used to extract text from images and then you can redact the text using the DLP API, but this process adds a layer of latency because it involves two steps. You can create custom infoTypes with regular expressions, but these are recommended only in situations where standard infoTypes aren't supported, such as medical account numbers.

D. Incorrect. The Cloud Vision API's OCR can be used to extract text from images and then you can redact the text using DLP API, but this process adds a layer of latency because it involves two steps. The Cloud Natural Language API cannot help with text redaction. Use the DLP API for this.

Where to look:
- https://cloud.google.com/dlp/docs/concepts-image-redaction
- https://cloud.google.com/dlp/docs/redacting-sensitive-data-images
- https://cloud.google.com/dlp/docs/infotypes-reference

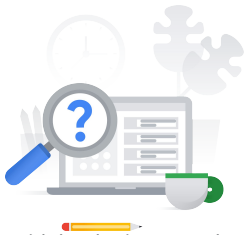
Content mapping:

- ILT course: **Security in Google Cloud**
  - M10 Content-related Vulnerabilities: Techniques and Best Practices

- On-demand course: **Mitigating Security Vulnerabilities on Google Cloud**
  - M2 Content-related Vulnerabilities: Techniques and Best Practices

Summary:
The Cloud Data Loss Prevention (DLP) API can redact images, documents, and text. It offers built-in and custom infoTypes that can help identify sensitive information. Using one of the supported programming languages, you can programmatically make redact image requests to the DLP API.

3.1 Diagnostic Question 02 Discussion



Cymbal Bank needs to statistically predict the days customers delay the payments for loan repayments and credit card repayments. Cymbal Bank does not want to share the exact dates a customer has defaulted or made a payment with data analysts. Additionally, you need to hide the customer name and the customer type, which could be corporate or retail.

How do you provide the appropriate information to the data analysts?

A. Generalize all dates to year and month with bucketing. Use the built-in infoType for customer name. Use a custom infoType for customer type with a custom dictionary.

B. Generalize all dates to year and month with bucketing. Use the built-in infoType for customer name. Use a custom infoType for customer type with regular expression.

C. Generalize all dates to year and month with date shifting. Use a predefined infoType for customer name. Use a custom infoType for customer type with a custom dictionary.

D. Generalize all dates to year and month with date shifting. Use a predefined infoType for customer name. Use a custom infoType for customer type with regular expression.

Google Cloud

Feedback:
A. Incorrect. Bucketing can reduce timestamps to smaller groups, but can also lose the sequence of events and time intervals. Generalizing all dates to year and month will also create difficulty for prediction, because the prediction must be in days. The usage of the built-in infoType for customer name and a custom infoType for customer type with a custom dictionary is correct, however.

B. Incorrect. Bucketing can reduce timestamps to smaller groups, but can also lose the sequence of events and time intervals. Generalizing all dates to year and month will also create difficulty for prediction, because the prediction must be in days. Usage of the built-in infoType for customer name is correct. Using a custom InfoType with regular expression will require additional filtering.

C. Correct! If your data is stored in a valid schema, date shifting will shift all dates logically. Built-in infoTypes allow a range of locale-specific and globally identifiable sensitive information pieces like email IDs and phone numbers. Custom dictionaries can be used with a custom infoType that contains predefined key-value pairs.

D. Incorrect. If your data is stored in a valid schema, date shifting will shift all dates logically. Usage of the built-in infoType for customer name is correct. Using a custom InfoType with regular expression will require additional filtering.

Where to look:
https://cloud.google.com/dlp/docs/concepts-date-shifting

https://cloud.google.com/dlp/docs/concepts-infotypes
https://cloud.google.com/dlp/docs/pseudonymization

Summary:
Date shifting and bucketing can help with date-time generalizing. Although bucketing can reduce the timestamp to smaller groups such as month or year, it can lose the sequence of events and time intervals. Date shifting shifts the date-time stamps while preserving the order and time intervals. The shifting period is different for different rows of data, so design the schema wisely.

Sensitive Data Protection uses built-in and custom infoType detectors to scan images, documents, and text. Custom infoTypes can be dictionaries, regular expressions, or dictionaries extracted from BigQuery or Cloud Storage. Use dictionaries when you want to match a list of words or phrases, and use regular expressions when you want to detect matches based on a regex pattern.

3.1 Diagnostic Question 03 Discussion

Cymbal Bank stores customer information in a BigQuery table called 'Information,' which belongs to the dataset 'Customers.' Various departments of Cymbal Bank, including loan, credit card, and trading, access the information table. Although the data source remains the same, each department needs to read and analyze separate customers and customer-attributes. You want a cost-effective way to configure departmental access to BigQuery to provide optimal performance.

What should you do?

A. Create separate datasets for each department. Create views for each dataset separately. Authorize these views to access the source dataset. Share the datasets with departments. Provide the bigquery.dataViewer role to each department's required users.

B. Create an authorized dataset in BigQuery's Explorer panel. Write Customers' table metadata into a JSON file, and edit the file to add each department's Project ID and Dataset ID. Provide the bigquery.user role to each department's required users.

C. Secure data with classification. Open the Data Catalog Taxonomies page in the Google Cloud Console. Create policy tags for required columns and rows. Provide the bigquery.user role to each department's required users. Provide policy tags access to each department separately.

D. Create separate datasets for each department. Create authorized functions in each dataset to perform required aggregations. Write transformed data to new tables for each department separately. Provide the bigquery.dataViewer role to each department's required users.

Google Cloud

Feedback:
A. Correct! Using authorized views is the right approach. Create a separate dataset for each department, and provide access to views containing filtered rows and columns.

B. Incorrect. There could be more tables or views in the 'Customers' dataset. Creating an authorized dataset will share all views inside it. The role bigquery.user is not sufficient because bigquery.dataViewer is required to query views.

C. Incorrect. Data classification could fit the scenario and add additional security on top of authorized views. However, authorized views have better performance and flexibility. The role bigquery.user is not sufficient because bigquery.dataViewer is required to query views.

D. Incorrect. Although authorized functions could fit the scenario with the help of user-defined functions (UDF), UDF execution is slower when compared to authorized views. Creating new tables would incur additional cost. Providing the bigquery.dataViewer role to each department's required users is correct.

Where to look:
- https://cloud.google.com/bigquery/docs/authorized-views
- https://cloud.google.com/bigquery/docs/authorized-datasets

Content mapping:

- ILT course: **Security in Google Cloud**
  - M6 Securing Cloud Data: Techniques and Best Practices

- On-demand course: **Security Best Practices in Google Cloud**
  - M2 Securing Cloud Data: Techniques and Best Practices

Summary:
Authorized views and data classification are two approaches for row-level security. Data classification helps with creating policies to access columns and rows. These policies can be assigned to users through Identity and Access Management (IAM). Authorized views help with providing limited control over the data. Both data classification and authorized views let you slice tables and provide different levels of access to BigQuery users.

3.1 Diagnostic Question 04 Discussion

Cymbal Bank has a Cloud SQL instance that must be shared with an external agency. The agency's developers will be assigned roles and permissions through a Google Group in Identity and Access Management (IAM). The external agency is on an annual contract and will require a connection string, username, and password to connect to the database.

How would you configure the group's access?

A. Use Secret Manager. Use the duration attribute to set the expiry period to one year. Add the secretmanager.secretAccessor role for the group that contains external developers.

B. Use Cloud Key Management Service. Use the destination IP address and Port attributes to provide access for developers at the external agency. Remove the IAM access after one year and rotate the shared keys. Add cloudkms.cryptoKeyEncryptorDecryptor role for the group that contains the external developers.

C. Use Secret Manager. Use the resource attribute to set a key-value pair with key as duration and values as expiry period one year from now. Add secretmanager.viewer role for the group that contains external developers.

D. Use Secret Manager for the connection string and username, and use Cloud Key Management Service for the password. Use tags to set the expiry period to the timestamp one year from now. Add secretmanager.secretVersionManager and secretmanager.secretAccessor roles for the group that contains external developers.

Google Cloud

Feedback:
A. Correct! Secret Manager supports time types such as absolute time duration to invoke and revoke access. The Secret Assessor role is required to read the stored secrets in Secret Manager.

B. Incorrect. You can use Cloud KMS to configure and manage Google-managed, customer-managed, and customer-supplied encryption keys. With Cloud KMS, developers will be able to decrypt shared information. Using IP address and Port ranges is incorrect because users are available as a group in IAM.

C. Incorrect. Secret Manager supports time types such as absolute time duration to invoke and revoke access. However, the Viewer role for Secret Manager does not allow users to use secrets.

D. Incorrect. You can use Cloud KMS to configure and manage Google-managed, customer-managed, and customer-supplied encryption keys. All details should be stored only in Secret Manager for our scenario. With Cloud KMS, developers will be able to decrypt shared information.

Where to look:
https://cloud.google.com/secret-manager/docs/access-control

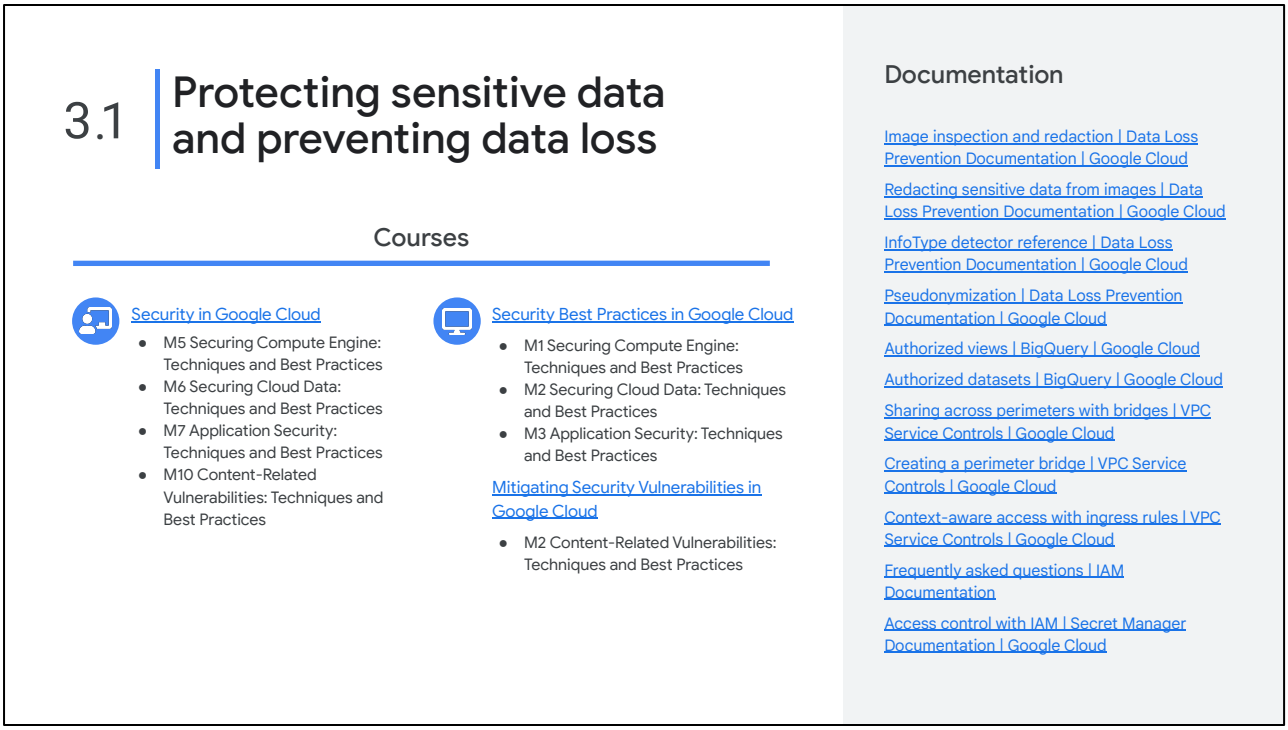
Content mapping:
- ILT course: **Security in Google Cloud**

    - M7 Application Security: Techniques and Best Practices

- On-demand course: **Security Best Practices in Google Cloud**
    - M3 Application Security: Techniques and Best Practices

Summary:
Secret Manager helps save confidential details such as passwords and URLs. You can provide access to secrets using IAM. Secret Manager lets organizations share configured secrets instead of confidential information with developers. Cloud KMS is used for storing encryption keys that are either managed by Google or by the customer.

Cloud KMS lets you share symmetric and asymmetric keys. Cloud KMS can be used to encrypt/decrypt data, but that will expose critical information to developers in plain text.



Let's take a moment to consider resources that can help you build your knowledge and skills in this area.

The concepts in the diagnostic questions we just reviewed are covered in these modules and in this documentation. Reviewing the documentation is highly recommended. You'll find this list in your workbook so you can take a note of what you want to include later when you build your study plan. Based on your experience with the diagnostic questions, you may want to include some or all of these.

- https://cloud.google.com/dlp/docs/concepts-image-redaction
- https://cloud.google.com/dlp/docs/redacting-sensitive-data-images
- https://cloud.google.com/dlp/docs/infotypes-reference
- https://cloud.google.com/dlp/docs/pseudonymization
- https://cloud.google.com/bigquery/docs/authorized-views
- https://cloud.google.com/bigquery/docs/authorized-datasets
- https://cloud.google.com/vpc-service-controls/docs/share-across-perimeters
- https://cloud.google.com/vpc-service-controls/docs/create-perimeter-bridges
- https://cloud.google.com/vpc-service-controls/docs/context-aware-access
- https://cloud.google.com/iam/docs/faq#how_do_i_grant_permissions_to_resources_in_my_project_to_someone_who_is_not_part_of_my_organization
- https://cloud.google.com/secret-manager/docs/access-control

3.2 Managing encryption at rest, in transit, and in use

Considerations include:

- Identifying use cases for Google default encryption, customer-managed encryption keys (CMEK), Cloud External Key Manager (EKM), and Cloud HSM
- Creating and managing encryption keys for CMEK and EKM
- Applying Google's encryption approach to use cases
- Configuring object lifecycle policies for Cloud Storage
- Enabling confidential computing

Google Cloud

A Professional Cloud Security Engineer should also have a clear understanding of the considerations involved in managing encryption at rest, and be able to apply Google's encryption approach to a variety of use cases. You should be familiar with customer-managed encryption keys (CMEK), Cloud External Key Manager (EKM), and Cloud HSM.

Question 5 tested your knowledge of Cloud Storage object lifecycle management. Question 6 asked you to identify the steps necessary to apply encryption settings. Question 7 tested your understanding of confidential VMs and their benefits.

3.2 Diagnostic Question 05 Discussion

Cymbal Bank calculates employee incentives on a monthly basis for the sales department and on a quarterly basis for the marketing department. The incentives are released with the next month's salary. Employee's performance documents are stored as spreadsheets, which are retained for at least one year for audit. You want to configure the most cost-effective storage for this scenario.

What should you do?

A. Import the spreadsheets to BigQuery, and create separate tables for Sales and Marketing. Set table expiry rules to 365 days for both tables. Create jobs scheduled to run every quarter for Marketing and every month for Sales.

B. Upload the spreadsheets to Cloud Storage. Select the Nearline storage class for the sales department and Coldline storage for the marketing department. Use object lifecycle management rules to set the storage class to Archival after 365 days. Process the data on BigQuery using jobs that run monthly for Sales and quarterly for Marketing.

C. Import the spreadsheets to Cloud SQL, and create separate tables for Sales and Marketing. For Table Expiration, set 365 days for both tables. Use stored procedures to calculate incentives. Use App Engine cron jobs to run stored procedures monthly for Sales and quarterly for Marketing.

D. Import the spreadsheets into Cloud Storage and create NoSQL tables. Use App Engine cron jobs to run monthly for Sales and quarterly for Marketing. Use a separate job to delete the data after 1 year.

Google Cloud

Feedback:
A. Incorrect. Although this solution works, it is not the most cost-effective. Use BigQuery if you need a high-performance solution to disburse salary immediately after processing. Use Cloud Storage and load external data into BigQuery to lower the cost.

B. Correct! Cloud Storage storage classes let you lower the storage cost for data that you access less frequently and don't require for real-time applications. Use object lifecycle rules to change storage classes and expire data. For processing, use BigQuery, which has a free daily quota.

C. Incorrect. Cloud SQL is a more expensive choice than Cloud Storage storage classes. App Engine has a free quota that can be used to run your cron jobs, but your solutions need to run once a month and once a quarter. App Engine is good for scenarios that require web availability or high availability.

D. Incorrect. Use Datastore for NoSQL transactional applications or semi-structured information such as categories, subcategories, product descriptions, logs, and variable sensor data. Using Datastore for highly structured information is possible, but is not the product use case. You will also need to run additional infrastructure to manipulate all aspects of data.

Where to look:
- https://cloud.google.com/storage/docs/storage-classes

- https://cloud.google.com/storage/docs/lifecycle

Content mapping:

- ILT course: **Security in Google Cloud**
  - M6 Securing Cloud Data: Techniques and Best Practices

- On-demand course: **Security Best Practices in Google Cloud**
  - M2 Securing Cloud Data: Techniques and Best Practices

Summary:

Cloud Storage lets you use storage classes that are less expensive than standard storage. These storage classes use low-performance HDDs compared to standard SDDs. The cost-performance trade-off lets you build low-cost, resilient applications that still have the fastest read access and lowest latency in the cloud space. Cloud Storage object lifecycle rules let you change storage classes or set expiration rules to further reduce storage costs.

3.2 Diagnostic Question 06 Discussion

Cymbal Bank uses Google Kubernetes Engine (GKE) to deploy its Docker containers. You want to encrypt the boot disk for a cluster running a custom image so that the key rotation is controlled by the Bank. GKE clusters will also generate up to 1024 randomized characters that will be used with the keys with Docker containers.

What steps would you take to apply the encryption settings with a dedicated hardware security layer?

A. In the Google Cloud console, navigate to Google Kubernetes Engine. Select your cluster and the boot node inside the cluster. Enable customer-managed encryption. Use Cloud HSM to generate random bytes and provide an additional layer of security.

B. Create a new GKE cluster with customer-managed encryption and HSM enabled. Deploy the containers to this cluster. Delete the old GKE cluster. Use Cloud HSM to generate random bytes and provide an additional layer of security.

C. Create a new key ring using Cloud Key Management Service. Extract this key to a certificate. Use the kubectl command to update the Kubernetes configuration. Validate using MAC digital signatures, and use a startup script to generate random bytes.

D. Create a new key ring using Cloud Key Management Service. Extract this key to a certificate. Use the Google Cloud console to update the Kubernetes configuration. Validate using MAC digital signatures, and use a startup script to generate random bytes.

Google Cloud

Feedback:
A. Incorrect. A Kubernetes cluster can be accessed through the kubectl command. Usage of Cloud HSM is correct.

B. Correct! Building a new cluster and deleting the old one is the solution. Cloud HSM provides an additional layer of dedicated hardware security and generates random bytes of up to 1024 characters.

C. Incorrect. Validating using MAC digital signatures is not helpful because they are used to verify messages. However, a startup script can be used to generate a random sequence.

D. Incorrect. The Google Cloud console cannot be used to edit a Kubernetes configuration. Validating using MAC digital signatures is not helpful because they are used to verify messages. However, a startup script can be used to generate a random sequence.

Where to look:
- https://cloud.google.com/kubernetes-engine/docs/how-to/using-cmek#boot-disks
- https://cloud.google.com/kubernetes-engine/docs/how-to/custom-boot-disks
- https://cloud.google.com/kms/docs/using-other-products#cmek_integrations

Content mapping:

- ILT course: **Security in Google Cloud**
  - M5 Securing Compute Engine: Techniques and Best Practices
  - M6 Securing Cloud Data: Techniques and Best Practices
  - M8 Securing Google Kubernetes Engine

- On-demand course: **Security Best Practices in Google Cloud**
  - M1 Securing Compute Engine: Techniques and Best Practices
  - M2 Securing Cloud Data: Techniques and Best Practices
  - M4 Securing Google Kubernetes Engine

- Skill badge: Implement Cloud Security Fundamentals on Google Cloud

Summary:
All Google Cloud resources default to have data encrypted at rest. You can use Google-managed encryption keys to further encrypt GKE clusters, Compute Engine instance boot disks, Cloud Storage, and BigQuery. Alternatively, you could use customer-managed encryption keys to rotate keys.

3.2 Diagnostic Question 7 Discussion

Cymbal Bank needs to migrate existing loan processing applications to Google Cloud. These applications transform confidential financial information. All the data should be encrypted at all stages, including sharing between sockets and RAM. An integrity test should also be performed every time these instances boot. You need to use Cymbal Bank's encryption keys to configure the Compute Engine instances.

What should you do?

A. Create a Confidential VM instance with Customer-Supplied Encryption Keys. In Cloud Logging, collect all logs for sevLaunchAttestationReportEvent.

B. Create a Shielded VM instance with Customer-Supplied Encryption Keys. In Cloud Logging, collect all logs for earlyBootReportEvent.

C. Create a Confidential VM instance with Customer-Managed Encryption Keys. In Cloud Logging, collect all logs for earlyBootReportEvent.

D. Create a Shielded VM instance with Customer-Managed Encryption Keys. In Cloud Logging, collect all logs for sevLaunchAttestationReportEvent.

Google Cloud

Feedback:
A. Correct! Use Customer-Supplied Encryption Keys because you need to use your own encryption keys. Confidential VMs have a unique launch attestation event that can be read from Cloud Logging.

B. Incorrect. Although using Customer-Supplied Encryption Keys is the correct choice, earlyBootReportEvent is not the parameter for booting. Use the launch attestation event from Cloud Logging.

C. Incorrect. Customer-Managed Encryption Keys will only let you manage key rotation, but you need to use your own encryption keys. Use the launch attestation event from Cloud Logging because earlyBootReportEvent is not the parameter for booting.

D. Incorrect. Customer-Managed Encryption Keys only let you manage key rotation, but you need to use your own encryption keys. However, Confidential VMs have a unique launch attestation event that can be read from Cloud Logging, so sevLaunchAttestationReportEvent is the correct choice.

Where to look:
https://cloud.google.com/compute/confidential-vm/docs/about-cvm

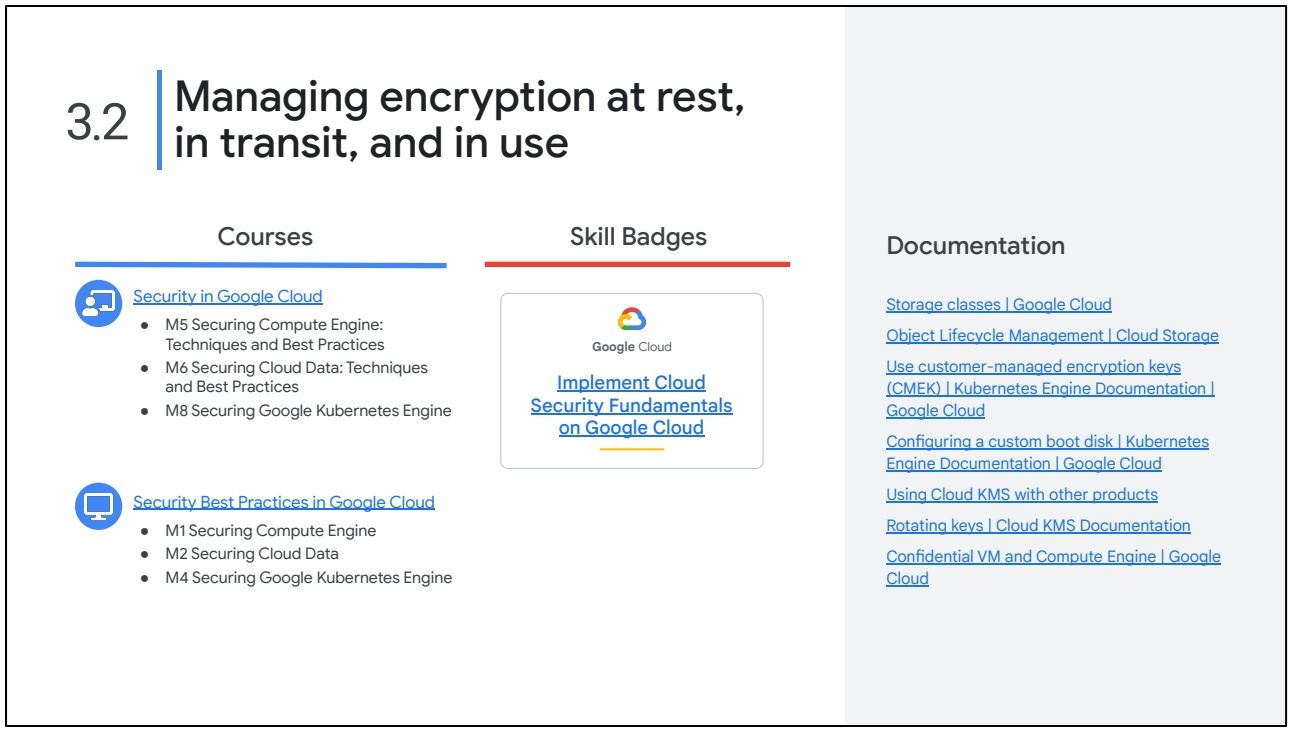
Content mapping:
- ILT course: **Security in Google Cloud**

    - M5 Securing Compute Engine: Techniques and Best Practices

- On-demand course: **Security Best Practices in Google Cloud**
    - M1 Securing Compute Engine: Techniques and Best Practices

Summary:
Confidential VMs use AMD's Secure Encrypted Virtualization, which keeps data encrypted in RAM. They can be managed using Customer-Supplied or Customer-Managed Encryption Keys. These instances contain dedicated AES engines that encrypt data as it flows out of sockets and decrypt data when it is read.

When restarting, Confidential VMs generate a unique log called Launch Attestation. Cloud Logging can be used to filter the logs and collect sevLaunchAttestationReportEvent.



Let's take a moment to consider resources that can help you build your knowledge and skills in this area.

The concepts in the diagnostic questions we just reviewed are covered in these modules and in this documentation. Reviewing the documentation is highly recommended. You'll find this list in your workbook so you can take a note of what you want to include later when you build your study plan. Based on your experience with the diagnostic questions, you may want to include some or all of these.

- https://cloud.google.com/storage/docs/storage-classes
- https://cloud.google.com/storage/docs/lifecycle
- https://cloud.google.com/kubernetes-engine/docs/how-to/using-cmek#boot-disks
- https://cloud.google.com/kubernetes-engine/docs/how-to/custom-boot-disks
- https://cloud.google.com/kms/docs/using-other-products#cmek_integrations
- https://cloud.google.com/kms/docs/rotating-keys
- https://cloud.google.com/compute/confidential-vm/docs/about-cvm

3.3 Planning for security and privacy in AI

Considerations include:

- Implementing security controls for AI/ML systems (e.g., protecting against unintentional exploitation of data or models)
- Determining security requirements for IaaS-hosted and PaaS-hosted training models

Google Cloud

A Professional Cloud Security Engineer should be able to implement security controls for AI/ML systems and to determine security requirements for IaaS-hosted and PaaS-hosted training models.

Question 8 tested your knowledge on how to protect data and prevent the misuse of a model. Question 9 asked you what should you prioritize when defining security requirements. Question 10 checked your knowledge on how to plan for AI/ML specific security controls when developing a system.

3.3 Diagnostic Question 08 Discussion

You are building an AI model on Google Cloud to analyze customer data and predict purchase behavior. This model will have access to sensitive information like purchase history and demographics.

A. Enable Google Cloud Armor on your deployed model to block malicious requests.
B. Store all model training data in BigQuery with public access for transparency.
C. Configure IAM roles to grant full access to the model for all Google Cloud users.
D. Deploy the model in a region with the highest data security standards.
E. Monitor the model's performance for anomalies and biases, then manually intervene if needed.

To protect this data and prevent misuse of the model, what THREE security controls are most important to implement?

Google Cloud

Feedback:
A. Correct! This actively protects the model from external threats and unauthorized access attempts, which is crucial for preventing data exploitation.

B. That's incorrect. While transparency is valuable, publicly exposing sensitive training data directly contradicts the goal of protecting it from unauthorized access and exploitation.

C. That's incorrect. This violates the principle of least privilege and significantly increases the risk of accidental or malicious misuse of the model.

D. Correct! This ensures compliance with regional regulations and data residency requirements, further safeguarding sensitive customer information.

E. Correct! Proactive monitoring and human oversight are essential for detecting potential misuse, bias, or unintended consequences of the AI model.

Where to look:
- https://cloud.google.com/blog/products/identity-security/how-sensitive-data-protection-can-help-secure-generative-ai-workloads

Content mapping:
- ILT course: **Security in Google Cloud**

    - M6 Securing Cloud Data
    - M10 Content-Related Vulnerabilities

- On-demand course: **Security Best Practices in Google Cloud**
    - M2 Securing Cloud Data: Techniques and Best Practices
- On-demand course: **Mitigating Security Vulnerabilities on Google Cloud**
    - M2 Content-Related Vulnerabilities: Techniques and Best Practices

Summary:
Leverage Google Cloud Armor, which provides a layer of protection against malicious requests targeting your deployed model. It acts as a firewall, filtering out traffic based on predefined rules and patterns designed to identify common web attacks.

Choose a deployment region that aligns with the highest data security standards applicable to your project. Consider factors like regulatory compliance (e.g., GDPR), industry-specific security certifications, and Google Cloud's own security measures in different regions.

Implement continuous monitoring of your model's performance. Look for anomalies that could indicate attempts to compromise the model's integrity or biases that may lead to discriminatory outcomes. Manual intervention allows you to take corrective actions when these issues are detected.

3.3 Diagnostic Question 09 Discussion

You're building a machine learning model on Google Cloud. You're choosing between two options: managing the infrastructure yourself (IaaS) or using Google's managed services (PaaS).

A. Network traffic inspection and intrusion detection
B. Compliance with internal security policies
C. Data location and residency restrictions
D. Granular access controls and permissions
E. Physical server hardening and security patches

To ensure the best security posture for both the model and its data, which TWO factors should you prioritize when defining security requirements for each hosting option?

Google Cloud

Feedback:

A. That's incorrect. While valuable for overall security, these measures are not specific to IaaS or PaaS hosting. Both options can implement them independently of the chosen hosting model.

B. That's incorrect. Internal security policies are important, but they apply equally to both IaaS and PaaS models and don't differentiate the specific security considerations for each hosting option.

C. Correct! In PaaS, Google manages the infrastructure, so you need to ensure data residency aligns with your compliance and privacy requirements.

D. Correct! Controlling who accesses the model and data is crucial for securing both IaaS and PaaS environments.

E. That's incorrect. While crucial for IaaS where you manage the underlying infrastructure, this becomes less relevant in PaaS. Google manages and secures the physical servers in PaaS, reducing your direct responsibility for hardening and patching.

Where to look:

- https://cloud.google.com/learn/paas-vs-iaas-vs-saas

Content mapping:

- ILT course: **Security in Google Cloud**
  - M2 Securing Access to Google Cloud
  - M10 Content-Related Vulnerabilities

- On-demand course: **Managing Security in Google Cloud**
  - M2 Securing Access to Google Cloud
- On-demand course: **Mitigating Security Vulnerabilities on Google Cloud**
  - M2 Content-Related Vulnerabilities: Techniques and Best Practices

Summary:

Understand the specific regulations and compliance requirements that dictate where your data can be stored and processed. Configure your cloud environment to enforce these restrictions to maintain control over data location.

Implement a fine-grained approach to authorization. Define precise permissions that determine who can access specific models and what actions they can perform. This minimizes the risk of unauthorized access or model misuse.

3.3 Diagnostic Question 10 Discussion

You are tasked with developing an AI system on Google Cloud for a telecommunications business. This AI system will conduct sentiment analysis on conversations agents have with customers, and provide conversational recommendations to improve customer satisfaction in the future.

What AI/ML-specific security controls do you need to plan for when developing this system?

A. Select Google Cloud AI services that leverage a PaaS model. These are the only ones that can guarantee a secure-by-design foundation.

B. Deploy your AI solution using managed instance groups (MIGs). These have baked in security controls specific to running AI workloads.

C. Leverage an AI model-specific threat detection scanner. Threats between AI systems and non-AI systems have very little in common.

D. AI systems are more interconnected than non-AI systems. Prepare for new attack vectors, as attackers can exploit vulnerabilities in one system to attack another.

Google Cloud

Feedback:
A. That's incorrect. All of Google Cloud's AI products are built atop a scalable technical infrastructure underpinned by a secure-by-design foundation and supported by robust logical, operational and physical controls to achieve defense in depth, at scale, and by default.

B. That's incorrect. MIGs do not contain security controls that are specific to running AI workloads.

C. That's incorrect. Many threats between AI systems and non-AI systems are the same. Both systems need to be protected from unauthorized access, modification, and destruction of data — as well as other common threats.

D. Correct! AI systems are more interconnected. AI systems are often connected to other systems, inside and outside of an organization. This interconnectedness can create new attack vectors, as attackers can exploit vulnerabilities in one system to attack another.

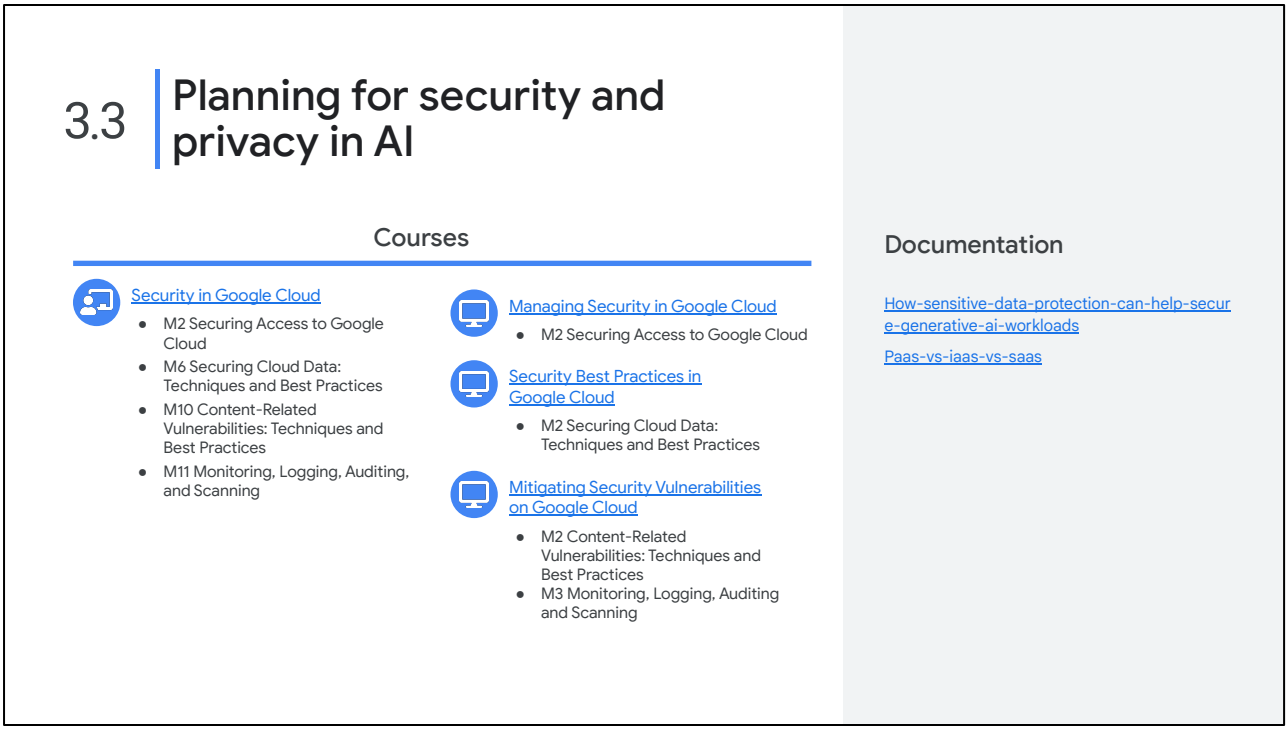
Content mapping:

- ILT course: **Security in Google Cloud**
  - M6 Securing Cloud Data
  - M11 Monitoring, Logging, Auditing, and Scanning

- On-demand course: **Security Best Practices in Google Cloud**

    - M2 Securing Cloud Data: Techniques and Best Practices
- On-demand course: **Mitigating Security Vulnerabilities in Google Cloud**
    - M3 Monitoring, Logging, Auditing and Scanning

Summary:
AI systems often have complex dependencies on other systems and data sources. This interconnectedness introduces new attack surfaces that may not be present in traditional software. Thoroughly map out these dependencies and potential vulnerabilities to proactively mitigate risks.



Let's take a moment to consider resources that can help you build your knowledge and skills in this area.

The concepts in the diagnostic questions we just reviewed are covered in these modules and in this documentation. Reviewing the documentation is highly recommended. You'll find this list in your workbook so you can take a note of what you want to include later when you build your study plan. Based on your experience with the diagnostic questions, you may want to include some or all of these.

- https://cloud.google.com/blog/products/identity-security/how-sensitive-data-protection-can-help-secure-generative-ai-workloads
- https://cloud.google.com/learn/paas-vs-iaas-vs-saas