Ensuring Solution Quality

Exam Guide

Designing for security and compliance.	• Identity and access management • Data security • Ensuring privacy • Legal compliance

Tip: IAM—Understand permissions and custom roles. Under what conditions are custom roles preferred over standard predefined roles?

Tip: Data security, data loss prevention—Cloud DLP allows you to minimize what you collect, store, expose, or copy. Classify or automatically redact sensitive data from text streams before you write to disk, generate logs, or perform analysis.

Be familiar with all these:

- Cloud IAM
- Encryption, Key Management
- Data Loss Prevention API
- HIPAA, COPPA, FedRAMP, GDPR

Exam Guide

Ensuring scalability and efficiency.	• Building and running test suites • Pipeline monitoring • Assessing, troubleshooting, and improving data representations and data processing infrastructure • Resizing and autoscaling resources

Tip: A lot of resource administration is presented in the GCP Console, but a lot of runtime information, such as logs and performance, is presented and reported in Stackdriver. Stackdriver provides information for troubleshooting both functional and performance issues.

- Stackdriver

Exam Guide

Ensuring reliability and fidelity.	• Performing data preparation and quality control • Verification and monitoring • Planning, executing, and stress testing data recovery • Choosing between ACID, idempotent, eventual consistency requirements

Tip: Establishing standard data quality at ingress by using Cloud Dataprep or by running an ETL pipeline can prevent many problems later in processing that would be difficult to troubleshoot.

Tip: Remember the business purpose of the data processing. How resilient does the application need to be? For example, financial transactions usually cannot be dropped and must not be duplicated, but a statistical analysis might be equally valid if a small amount of data is lost. These assumptions influence the approach to rerunning failed jobs.

Study these:

- Cloud Dataprep
- Fault-tolerance
- Rerunning failed jobs
- Performing retrospective re-analysis

Exam Guide

Ensuring flexibility and portability.	• Mapping to current and future business requirements • Designing for data and application portability

Tip: Where is the official authoritative data (sometimes called the source of truth) and where are the replicas? How frequently does data need to be shared or updated? Can smaller parts of the data be synchronized to reduce costs?

Tip: Where is the data stored? Where is the data going to be processed? Can data storage and data processing be in locations near each other?

Tip: When will the data need to be exported? How difficult and expensive will it be to export data? For example, you might want to store data in a different location or in a different type of storage to meet business requirements for portability.

- Multi-cloud data residency requirements