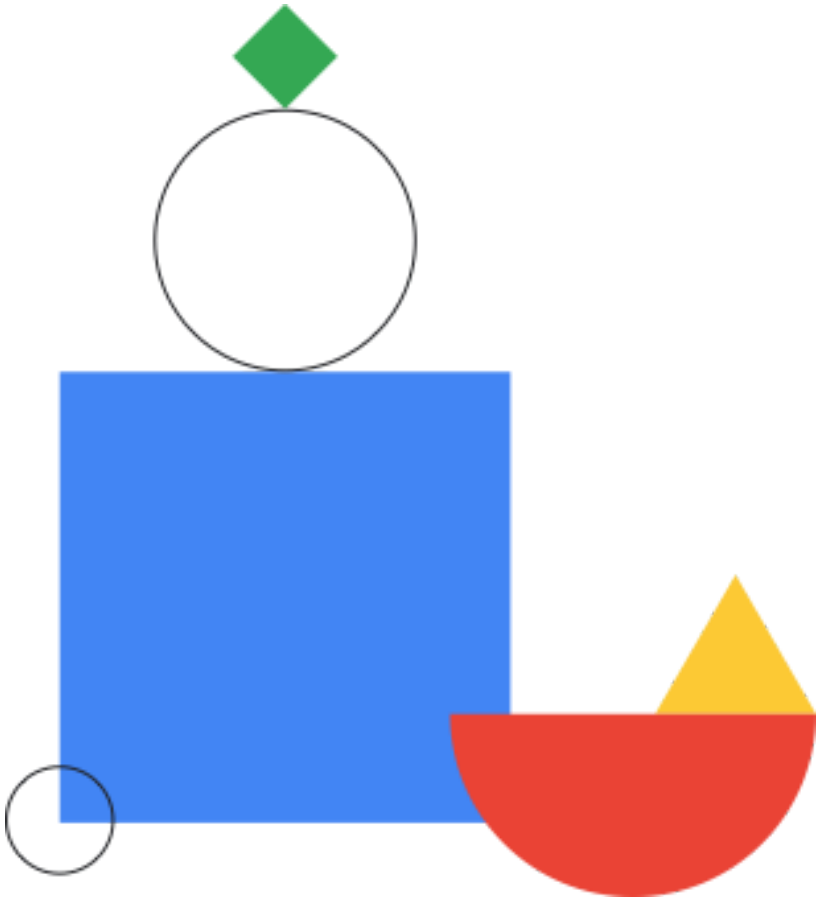
Google Cloud

Cloud Digital Leader Study Guide

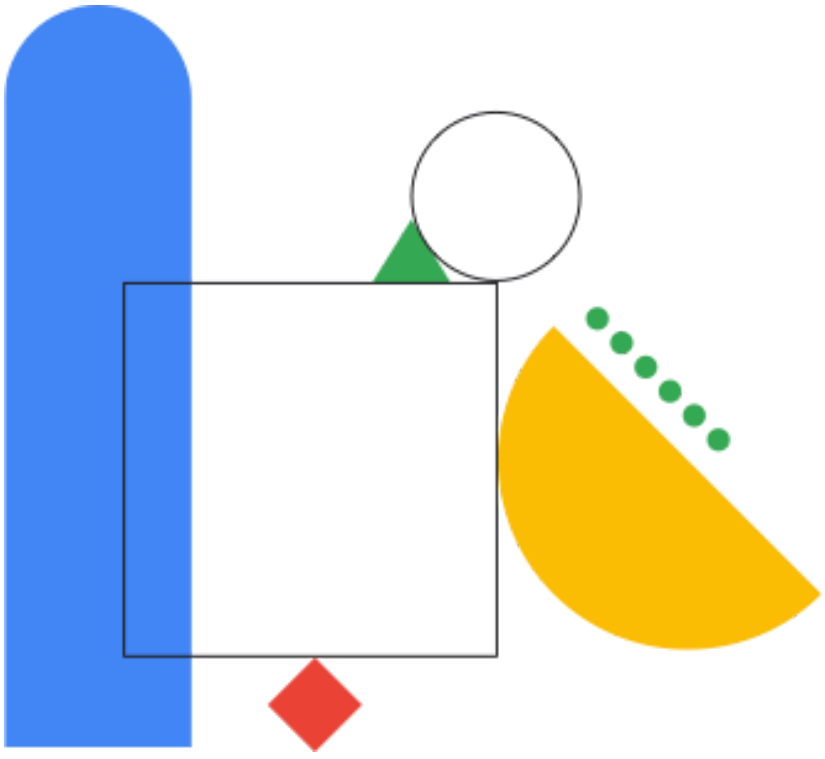


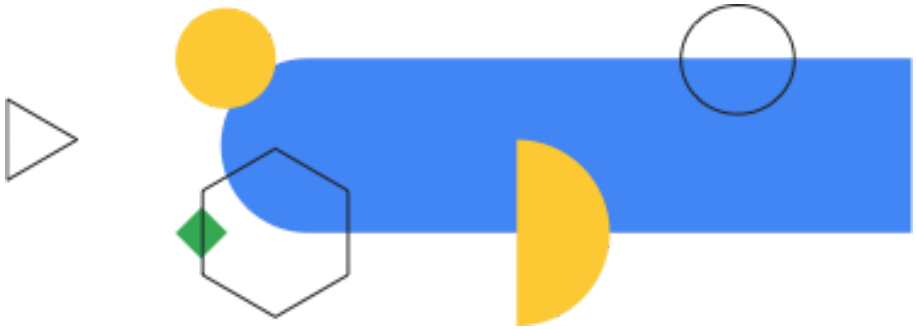
Introduction

The Google Cloud Digital Leader training and exam are intended for tech-adjacent individuals who want to demonstrate an overall knowledge of cloud technology concepts and Google Cloud.

The exam validates a candidate's ability to complete the following course objectives:

- Recall fundamental cloud terminology.
- Identify Google Cloud products and solutions that support digital transformation.
- Explain how cloud technology and data can be leveraged to innovate within organizations.
- Explain infrastructure and application modernization with Google Cloud.
- Describe contributing factors to the success of cloud operations and security.





Learn more about the exam

A Cloud Digital Leader can distinguish and evaluate the various capabilities of Google Cloud core products and services and how they can be used to achieve desired business goals. A Cloud Digital Leader is well-versed in fundamental cloud concepts and can demonstrate a broad application of cloud computing knowledge in a variety of applications.

The Cloud Digital Leader exam is job-role independent. The exam assesses the knowledge and skills of any individuals who want (or are required to) understand the purpose and application of Google Cloud products.

The Cloud Digital Leader exam assesses knowledge in four areas:

- General cloud knowledge (approximately 10% of the exam)
- Innovating with data and Google Cloud (approximately 30% of the exam)
- Modernizing infrastructure and applications with Google Cloud (approximately 30% of the exam)
- Understanding Google Cloud security and operations (approximately 30% of the exam)

Sign up for the Cloud Digital Leader Learning Path via Google Cloud Skills Boost, Coursera, or Pluralsight.

Prepare for the exam with sample questions.

Learn more about how and where to take the exam on the Cloud Digital Leader website.

Outline of learning path content

For learners who have taken Google Cloud's instructor-led training the terminology differs slightly. Each course listed here refers to one module and each module refers to a lesson.

Course 1

Digital Transformation with Google Cloud

Module 1: Why cloud technology is revolutionizing business

Module 2: Fundamental cloud concepts

Module 3: Cloud computing models and shared responsibility

Course 2

Innovating with Data and Google Cloud

Module 1: The value of data

Module 2: Data consolidation and analytics

Module 3: Innovation with machine learning

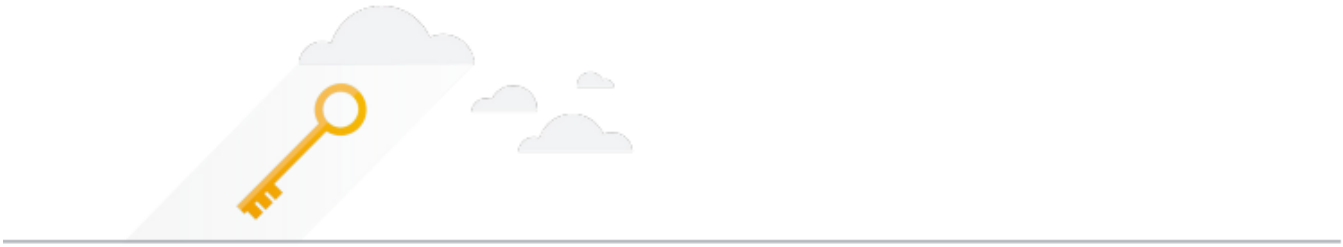
Course 3

Infrastructure and Application Modernization with Google Cloud

Module 1: Modernizing IT infrastructure with Google Cloud

Module 2: Modernizing applications with Google Cloud

Module 3: The value of APIs



Course 4

Understanding Google Cloud Security and Operations

Module 1: Financial governance in the cloud

Module 2: Security in the cloud

Module 3: Monitoring cloud IT services and operations

Course 1: Digital Transformation with Google Cloud

Module 1: Why cloud technology is revolutionizing business

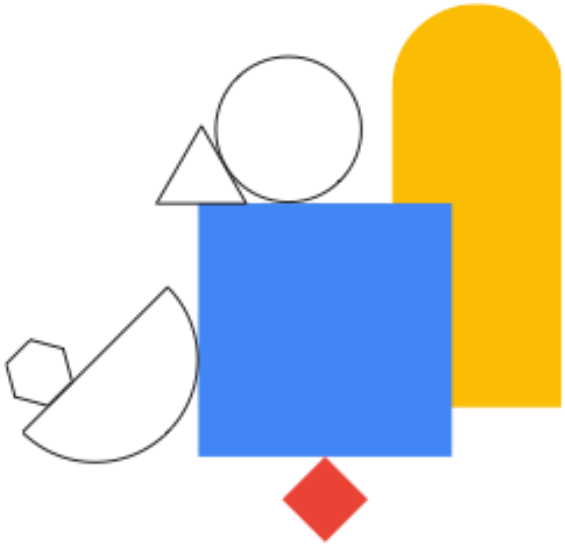
This module covers innovations throughout history that led to paradigm shifts, and looks at how cloud technology is enabling a new wave of digital transformation. It defines some important terms that you'll hear throughout the course, and it describes the benefits of adopting cloud technology to digitally transform a business. It also provides real-world examples and discusses the challenges that lead to a digital transformation.

Module 2: Fundamental cloud concepts

This module explains some fundamental cloud concepts and explains how migrating to the cloud affects your organization's flexibility, agility, reliability, and total cost of ownership. It explores the different types of infrastructure and various use cases for them. It also describes the Google Cloud network, how it's structured, and how performance is measured and impacted.

Module 3: Cloud computing models and shared responsibility

This module explains the different cloud computing models: IaaS (infrastructure as a service), PaaS (platform as a service), and SaaS (software as a service). It identifies the benefits and tradeoffs of implementing each computing model and how to choose the most appropriate one to meet business needs. It also explores the shared responsibility between an organization and its cloud provider in hardware, software, and security.



Key terms:

Cloud, cloud technology, compute power, computing, data, innovation

Focus areas for the exam:

- The difference between cloud technology and traditional or on-premises technology.
- How customer expectations have changed because of cloud technology.
- The business and technical considerations that organizations need to think about when adopting cloud technology.

Additional reading

What is cloud computing?

What is digital transformation?

Cloud computing 101: Frequently asked questions

What is Infrastructure as a Service (IaaS)?

Google Cloud products list page

Google's Guide to Innovation

Google Cloud Adoption Framework Impact Study

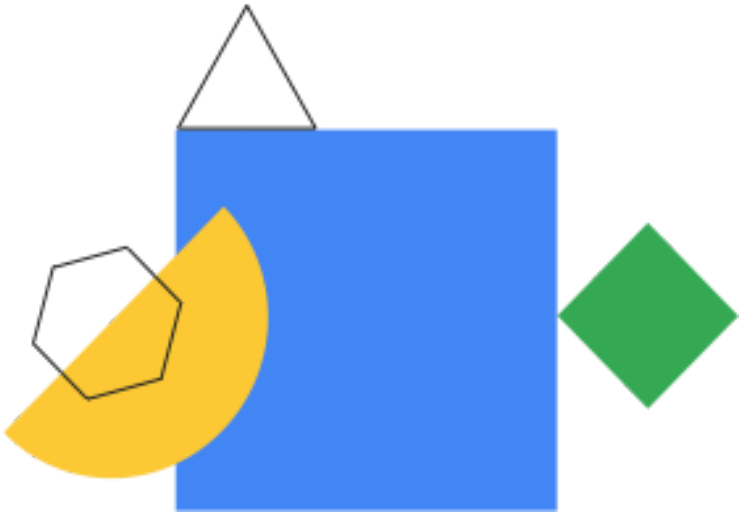
Managing Change in the Cloud

Cloud Customer Case Studies

Google Cloud whitepapers

Cloud locations

Google Cloud infrastructure (interactive web tool)



Course 2: Innovating with Data and Google Cloud

Module 1: The value of data

This module answers key questions about the nature of data and its role in a business' digital transformation. It discusses how data can be leveraged in an organization, breaks down the different types of data, and covers some important data considerations for every business using data in the cloud.

Module 2: Data consolidation and analytics

This module compares traditional data management methods with the benefits of migrating data to the Cloud. It defines key terms related to data management and covers some use cases and applicable Google Cloud solutions. It also explores business intelligence solutions that enable businesses to gain insight into their data.

Module 3: Innovation with machine learning

This module defines machine learning (ML) and artificial intelligence (AI). It covers some important data quality considerations that influence the efficacy of ML models. It also highlights several real-world use cases in which customers have leveraged ML to radically transform their business.

Key terms:

Data, data management, structured data, unstructured data, semi-structured data, databases, data warehouses, data lakes, artificial intelligence (AI), machine learning (ML), data coverage, data cleanliness/consistency, data completeness

Focus areas for the exam:

- How cloud technology enables data to be applied in new ways
- The benefits of storing and managing data in the cloud
- Appropriate use cases for data solutions
- How Google Cloud data products create business insights
- The importance of data quality in ML

Additional reading

What is a data lake?

What is a data warehouse?

What is streaming analytics?

What is business intelligence?

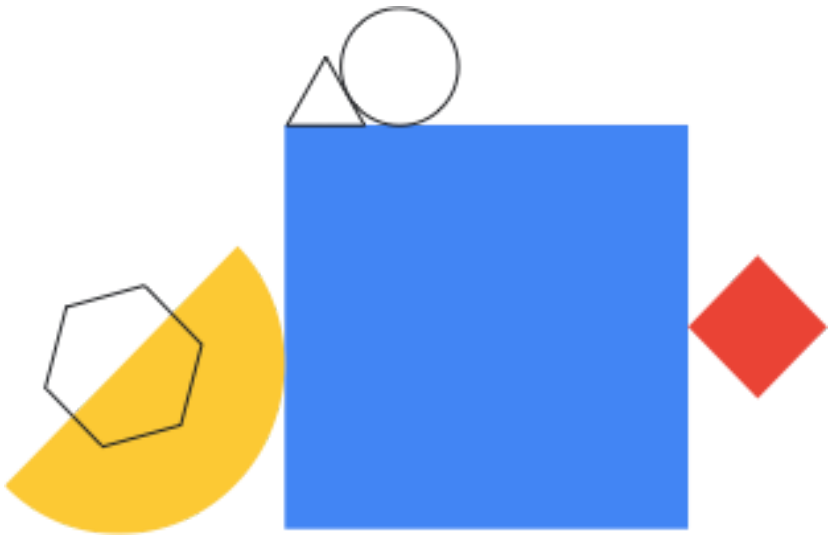
Build a modern, unified data analytics platform with Google Cloud

Principles and best practices for data governance in the cloud

What is machine learning?

Google Cloud's AI Adoption Framework

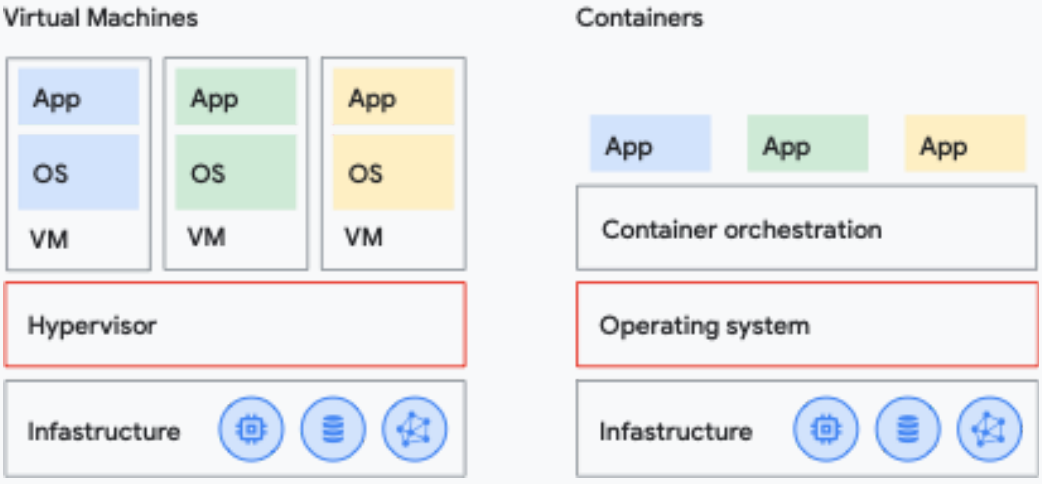
Google Cloud Responsible AI



Course 3: Infrastructure and Application Modernization with Google Cloud

Module 1: The value of data

This module answers key questions about the nature of data and its role in a business' digital transformation. It discusses how data can be leveraged in an organization, breaks down the different types of data, and covers some important data considerations for every business using data in the cloud.

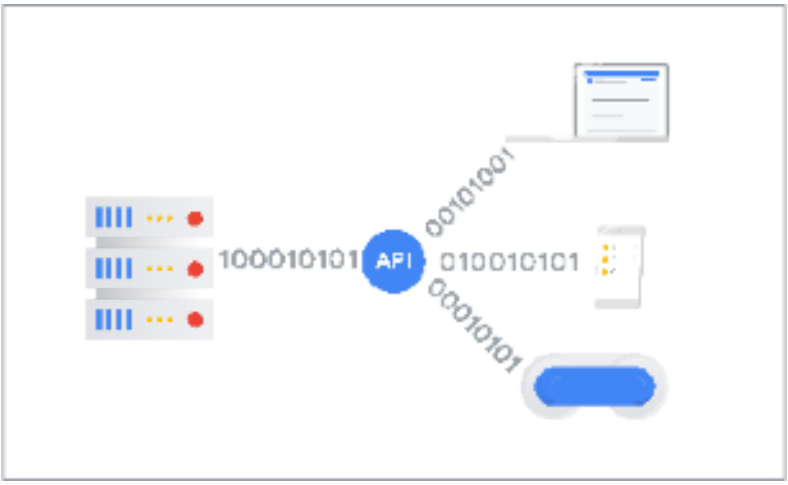


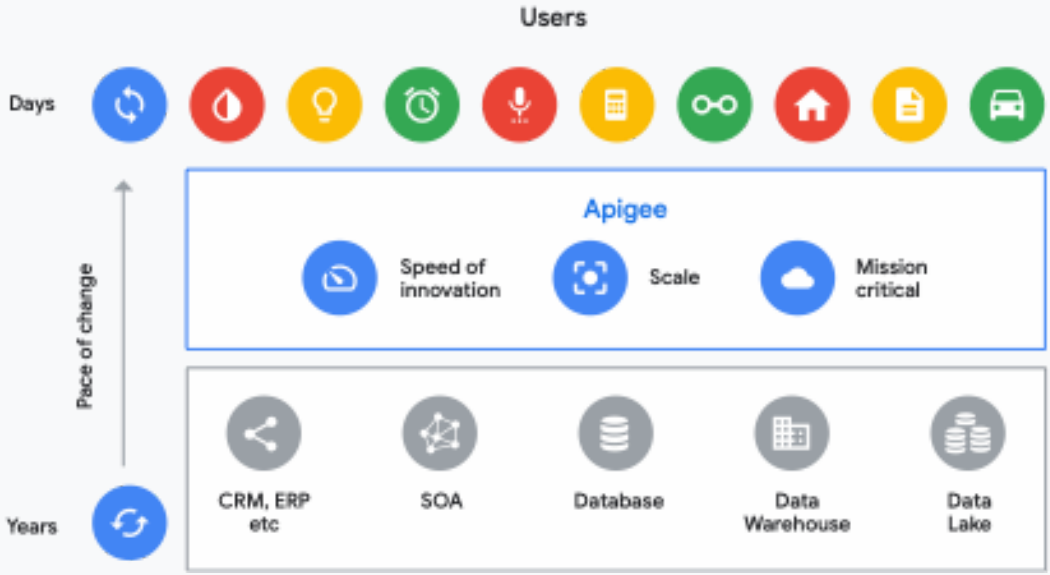
Module 2: Modernizing applications with Google Cloud

This module explores application modernization and common patterns of moving applications to the cloud such as the greenfield strategy. Multiple examples demonstrate how cloud technologies such as App Engine enables businesses to develop, deploy, and update applications with speed, security, and agility.

Module 3: The value of APIs

This module covers Application Programming Interfaces (APIs) and how they unlock value from legacy systems, enabling businesses to create new value by making more services to connect to customers. It also explores how Apigee and API management systems address application development gaps.





Key terms:

Infrastructure modernization, hybrid cloud, multi-cloud, virtual machines, containers, serverless commuting, applications, APIs

Focus areas for the exam:

- The benefits of modernizing infrastructure with cloud technology
- The difference between different architectures and compute options
- Google Cloud solutions for infrastructure modernization and application modernization
- The benefits of cloud-native applications and change patterns
- How APIs modernize legacy systems and create new business value
- The benefits of Apigee

Additional reading

Google Cloud overview

Google Cloud geography and regions

Best practices for enterprise organizations

Google Cloud setup checklist

What are containers?

What is hybrid cloud?

What is a virtual machine?

What is object storage?

Migrate workloads to the public cloud: an essential guide & checklist

Where should I run my stuff? Choosing a Google Cloud compute option

CIO's Guide to Application Migration

CIO's Guide to Application Modernization

Google Cloud Networking overview blog

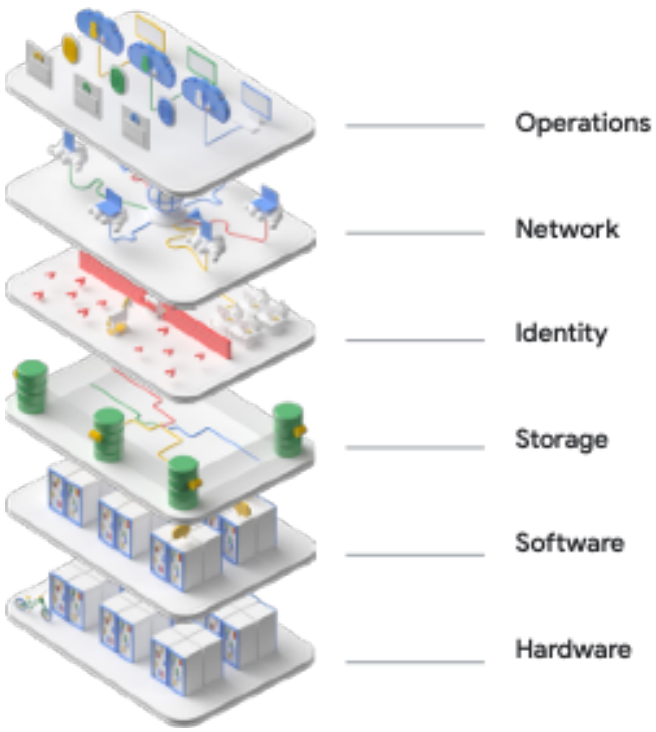
Course 4: Understanding Google Cloud Security and Operations

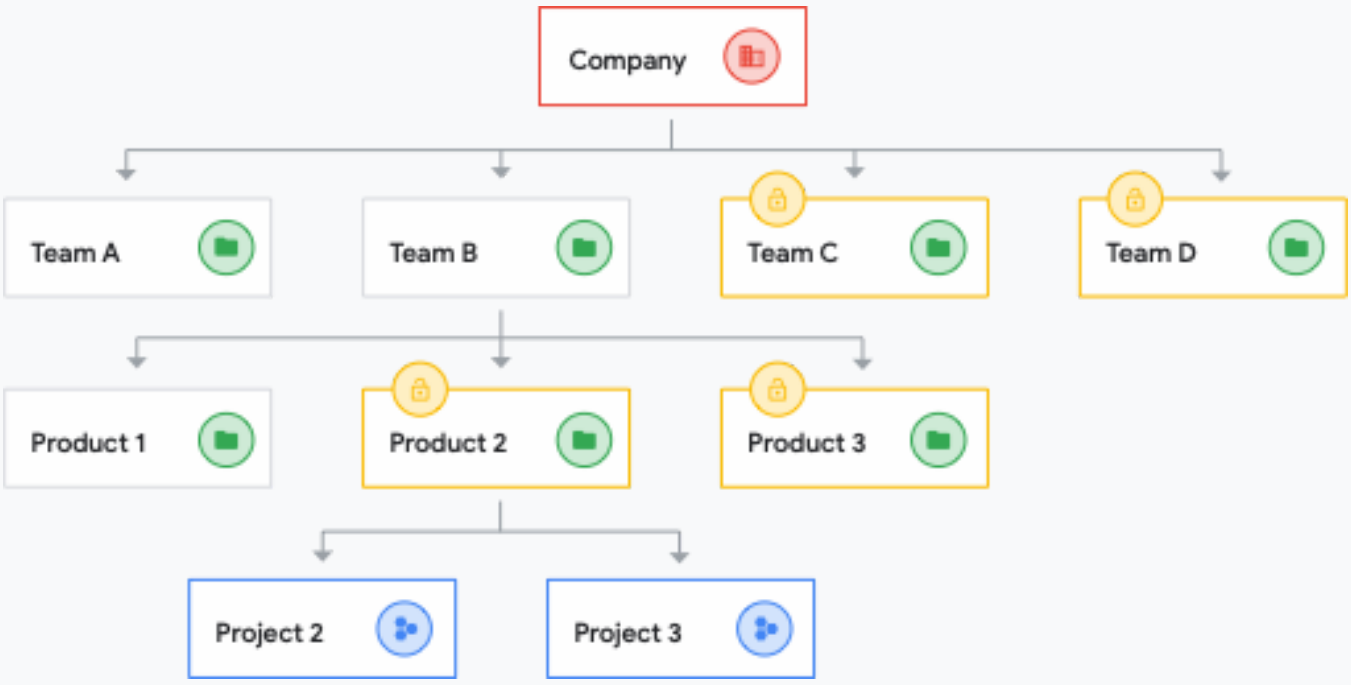
Module 1: Financial governance in the cloud

This module covers the fundamentals of cloud cost management and outlines how businesses can choose to maintain some (or none) of their own infrastructure by purchasing IT services from a cloud provider on an ongoing basis. It also explores the best practices for managing Google Cloud costs.

Module 2: Security in the cloud

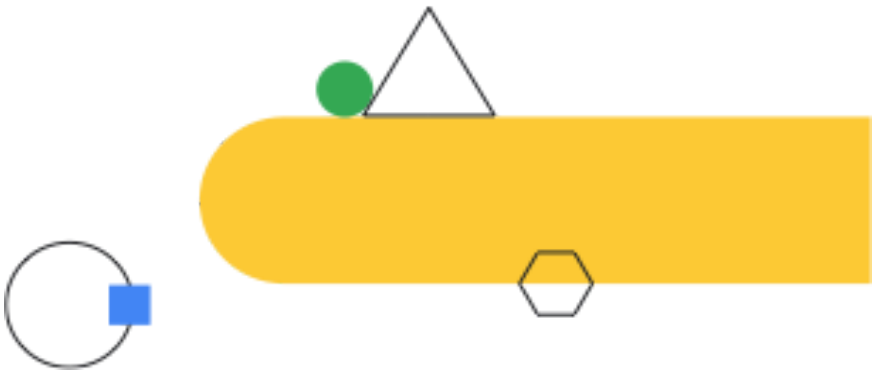
This module examines the shared responsibility model, which refers to how a business that accesses, stores or uses data is responsible for protecting and securing that data alongside the cloud provider. It also looks at the top cyber security threats and the defense-in-depth security built into Google Cloud.

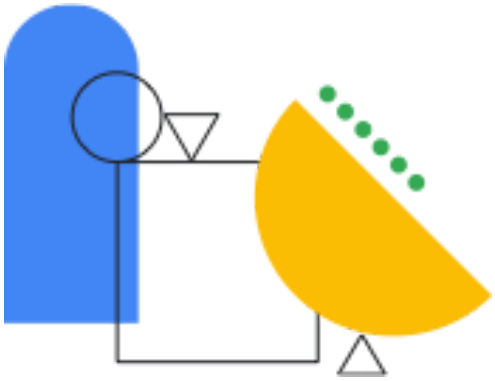




Module 3: Monitoring cloud IT services and operations

This module discusses how IT teams and business leaders need to rethink IT resource management in the cloud. This includes adjusting expectations to balance service availability and reliability with agile application development, as well as creating a team structure to better serve customers. Finally, the module covers how Google Cloud resource monitoring tools such as Cloud Identity, Cloud Debugger and Cloud Logging can help them to maintain visibility, accountability, control, and intelligence over their cloud environment.





Key terms:

Total Cost of Ownership (TCO), visibility, accountability, control, intelligence, privacy, availability, security, control, shared responsibility model, resource hierarchy,
SLO, SLA, SLI, DevOps, SRE, downtime, monitoring, logging, phishing attack, physical damage, malware and viruses, unsecured third party systems, expert knowledge, data at rest, data in transit, encryption at rest, encryption in transit, zero trust model, least privileged model

Focus areas for the exam:

- How cloud technology impacts TCO
- Cost management strategies
- Fundamental cloud security terms and top cybersecurity threats
- Security with Google Cloud
- How IT operations need to change in the cloud, including service availability expectations, DevOps and SRE
- Cloud resource monitoring and application performance management

Additional reading

Google Cloud security foundations guide

Google security

A guide to financial governance in the cloud

Google Infrastructure Security Design Overview

Increasing business value with better IT operations: A guide to site reliability engineering (SRE)

Exploring container security: the shared responsibility model in GKE

Resource hierarchy

AirAsia: adopting a modern identity solution with Google Cloud

Voyage to the modern workplace

Encryption at rest in Google Cloud

Encryption in transit in Google Cloud

Glossary

Course 1

The cloud – A metaphor for the network of data centers that store and compute information available through the Internet. It includes the complex web of software, computers, networks and security systems involved.

Cloud technology/computing – The technology and processes needed to store, manage, and access data that is transferred over the Cloud (as opposed to data that remains on your computer's hard drive).

Digital transformation – When an organization takes advantage of new technologies to redesign and redefine relationships with their customers, employees, and partners. Digital transformation uses modern digital technologies—including all types of public, private, and hybrid cloud platforms—to create or modify business processes, culture, and customer experiences to meet changing business and market dynamics.

Computing – A machine's ability to process, store, retrieve, compare and analyze information, and to automate tasks often done by computer programs (otherwise known as software or applications). Compute power – The speed at which a computer can process data.

Data – Any information that is useful to an organization. Can be numbers on a spreadsheet, text in an email, audio or video recordings, images, or even ideas in employees' heads. Includes internal and external information.

On-premises IT infrastructure – Often abbreviated to "on-prem," this refers to hardware and software applications that are hosted on-site, located and operated within an organization's data center to serve their unique needs.

Private cloud – When an organization has virtualized servers in its own data centers, or those of a private cloud provider, to create its own private dedicated environment.

Public cloud – Where on-demand computing services and infrastructure are managed by a third-party provider, such as Google Cloud, and shared with multiple organizations or "tenants" through the public internet.

Hybrid cloud – When applications run in a combination of private and public clouds.

Multicloud – When applications run in a combination of at least two public cloud providers.

Open standard – Software that follows particular specifications that are openly accessible and usable by anyone.

Open source – Software with source code that is publicly accessible and free for anyone to use, modify, and share.

Capital expenditures (CapEx) – Upfront business expenses put toward fixed assets. Organizations buy these items once, and they benefit their business for many years.

Operating expenses (OpEx) – Recurring costs for a more immediate benefit. This represents the day-to-day expenses to run a business.

Bandwidth – A measure of how much data a network can transfer in a given amount of time.

Network latency – The amount of time it takes for data to travel from one point to another. Often measured in milliseconds, latency, sometimes called lag, describes delays in communication over a network.

Zone – A geographic area where Google Cloud resources are deployed.

Regions – Independent geographic areas where Google Cloud resources are deployed, composed of zones.

Infrastructure as a service (IaaS) – A computing model that offers the on-demand availability of almost infinitely scalable infrastructure resources, such as compute, networking, storage, and databases as services over the internet.

Platform as a service (PaaS) – A computing model that offers a cloud-based platform for developing, running, and managing applications.

Software as a service (SaaS) – A computing model that offers an entire application, managed by a cloud provider, through a web browser

Shared responsibility model – A model in which the responsibility to secure data is shared between a business and the cloud provider. The cloud service provider is the data processor, while the organization is the data controller.

Course 2

Data point – A piece of information (e.g. a customer purchase or return).

Dataset – Aggregated data points of one category (e.g. customer transactions).

Structured data – Highly organized, quantitative data (e.g. names or credit card numbers). Easily stored and managed in databases.

Unstructured data – Data that has no organization and tends to be qualitative (e.g. word processing documents or images). Can be stored as objects, which consist of the data in its native format along with metadata such as unique identifiers.

Semi-structured data – Data that falls somewhere in between structured and unstructured data. It's organized into a hierarchy, but without full differentiation or any particular ordering. Examples include emails, HTML, JSON, and XML files.

Object storage – A data storage architecture for large stores of unstructured data, designating each piece of data as an object (e.g. audio or multimedia files).

Database – An organized collection of data generally stored in tables and accessed electronically from a computer system. Built and optimized to enable the efficient ingestion of large amounts of data from many different sources.

Data integrity – Also known as transaction integrity, this refers to the accuracy and consistency of data stored in a database.

Metadata – Information about objects (e.g. about images or audio).

Data warehouse – The central hub for all business data, it assembles data from multiple sources, including databases. When combined with connector tools, it can transform unstructured data into semi-structured data that can be used for analysis. Data warehouses are built to rapidly analyze and report massive and multi-dimensional datasets on an ongoing basis, in real-time.

Data lake – A data lake is a repository designed to store, process, and secure large amounts of structured, semistructured, and unstructured data. It can store data in its native format and process any variety of it, ignoring size limits and serves many purposes, such as exploratory data analysis.

Artificial intelligence (AI) – A broad field or term that describes any kind of machine capable of a task that normally requires human intelligence, such as visual perception, speech recognition,decision-making, or translation between languages.

Machine learning (ML) – A branch within the field of AI. Computers that can "learn" from data and make predictions or decisions without being explicitly programmed to do so, using algorithms or models to analyze data. These algorithms use historical data as input to predict new output values.

Course 3

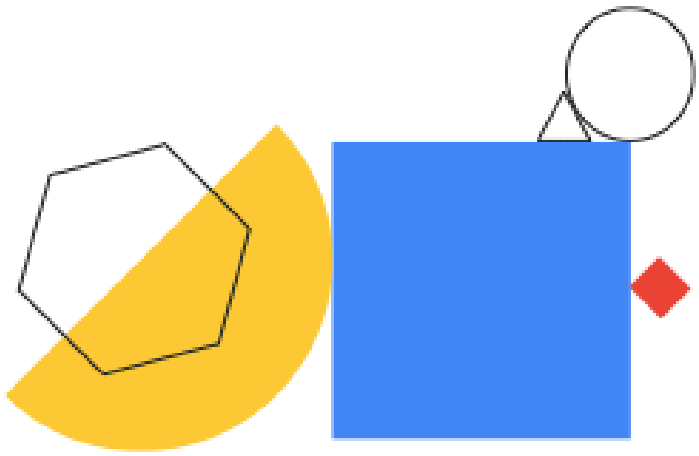
Virtual machines (VM) – A VM is a virtualized instance of a server that re-creates the functionality of a dedicated physical server. It uses a partitioned space inside a physical server which makes it easy to optimize and reallocate resources and allow multiple systems to run on the same hardware.

Hypervisor – The software layer that sits on top of physical hardware. Multiple VMs are built on top of the hypervisor and are enabled by it.

Container – Follows the same principle as a VM, providing an isolated environment to run software services and optimize resources from one piece of hardware. Containers are more efficient than VMs because they do not recreate a full representation of the hardware, but only recreate or virtualize the operating system.

Kubernetes – An open source cluster management system that provides automated container orchestration.

Serverless computing – A cloud computing execution model in which the cloud provider allocates machine resources on demand, taking care of the servers on behalf of their customers. Businesses provide code for the function they want to run and the cloud provider handles all infrastructure management. Resources such as compute power are automatically provisioned behind the scenes as needed.



Application (or app) – A computer program or software that is designed to carry out a specific digital task, typically used or run by an end-user. In this digital age, customers expect applications to be intuitive, well-functioning, and efficient.

Application programming interface (API) – A piece of software that interfaces with or connects different applications and enables information to flow between systems. In contrast to a user interface, which connects a computer to a person, an API connects computers or pieces of software to each other. One purpose of APIs is to hide the internal details of how a system works, exposing only those parts a developer wants to allow a user or program to interface with. In this way APIs can help organizations to adapt to modern business needs by allowing access to older legacy systems.

Course 4

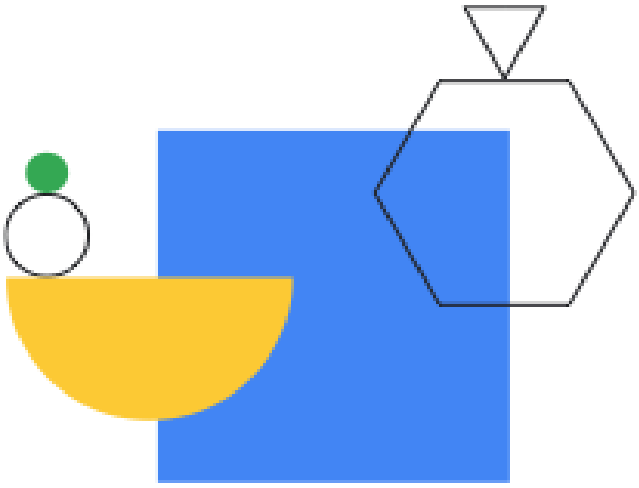
Total cost of ownership (TCO) – A comprehensive assessment of all layers within the infrastructure and other associated costs across the business over time. Includes acquiring hardware and software, management and support, communications, and user expenses, and the cost of service downtime, training and other productivity losses.

Privacy – The data an organization or an individual has access to, and who they can share that data with.

Security – The policies, procedures and controls put in place to keep data and infrastructure safe.

Availability – The duration for which the cloud service provider guarantees that client's data and services are up and running or accessible.

Defense-in-depth – The cloud service provider manages the security of its infrastructure and its data centers, and customers gain the benefits of their infrastructure's multiple built in security layers.



Resource hierarchy – How an IT team can organize a business's Google Cloud environment and how that service structure maps to the organization's actual structure. It determines what resources users can access.

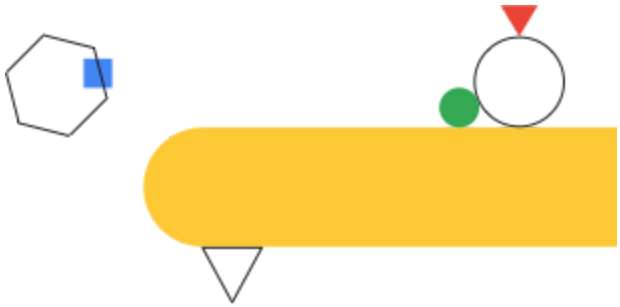
DevOps – Developer operations. A philosophy that seeks to create a more collaborative and accountable culture within developer and operations teams. Five objectives of DevOps include reducing silos, accepting failure as normal, implementing gradual change, leveraging tooling and automation and measuring everything.

SRE – Site reliability engineering. A discipline that applies aspects of software engineering to operations. The goals of SRE are to create ultra-scalable and highly reliable software systems. Best practices central to SRE align with DevOps objectives.

Monitoring – Gathering predefined sets of metrics or logs. Monitoring is the foundation for site reliability engineering because it provides visibility into the performance, uptime, and overall health of cloud powered applications.

Log file – A text file where applications (including the operating system) write events. Log files make it easier for developers, DevOps and system administrators to get insights and identify the root cause of issues within applications and the infrastructure.

Logging – A process that allows IT teams to analyze selected logs and accelerate application troubleshooting.



List of Google products and solutions

For more information about Google Cloud products

For more information about Google Cloud Solutions

Blog: 21 Google Cloud tools, each explained in under 2 minutes

Apigee API Management – A platform for developing and managing APIs.

App Engine – A platform for building scalable web applications and mobile back-ends.

Bare Metal – Infrastructure to run specialized workloads on Google Cloud.

BigQuery – Google Cloud's leading data warehouse solution.

Cloud Bigtable – Google's NoSQL big data database service.

Cloud Console – A web-based interface for managing and monitoring cloud apps.

Cloud Debugger – A real-time application state inspection and in-production debugging tool.

Cloud Functions – An event-driven compute platform for cloud services and apps.

Cloud Identity – A unified platform for IT administrators to manage user devices and apps.

Cloud Logging – An audit, platform, and application logs management tool.

Cloud Monitoring – A tool monitoring infrastructure and application health with rich metrics.

Cloud Profiler – Continuous CPU and heap profiling to improve performance and reduce costs.

Cloud Run – A fully managed environment for running containerized apps.

Cloud Spanner – A fully managed Google Cloud database service designed for global scale.

Cloud SQL – Google Cloud's database service (relational database management service).

Cloud Storage – Google Cloud's object storage service for structured, semi-structured and structured data. One of several products used in data lake solutions.

Cloud Trace – A tracing system collecting latency data from applications.

Compute Engine – Virtual machines running in Google's data center.

Cost Management – Tools for monitoring, controlling, and optimizing business costs.

Dataflow – A fully managed streaming analytics service that creates a pipeline to process both streaming data and batch data.

Firebase – An app development software to build, improve, and grow mobile and web apps.

Google Kubernetes Engine – An open-source container orchestration system for automating computer application deployment, scaling, and management.

Looker – Google Cloud's business intelligence solution.

Pub/Sub – A distributed messaging service that can receive messages from various device streams such as gaming events, IoT devices, and application streams. The name is short for Publisher/Subscriber.

TensorFlow – An end-to-end open source platform for machine learning, with a comprehensive, flexible ecosystem of tools, libraries and community resources, originally created by Google.

Vertex AI – A unified platform for training, hosting and managing ML models. Features include AutoML and custom training.

VMware Engine – An engine for migrating and running VMware workloads natively on Google Cloud.