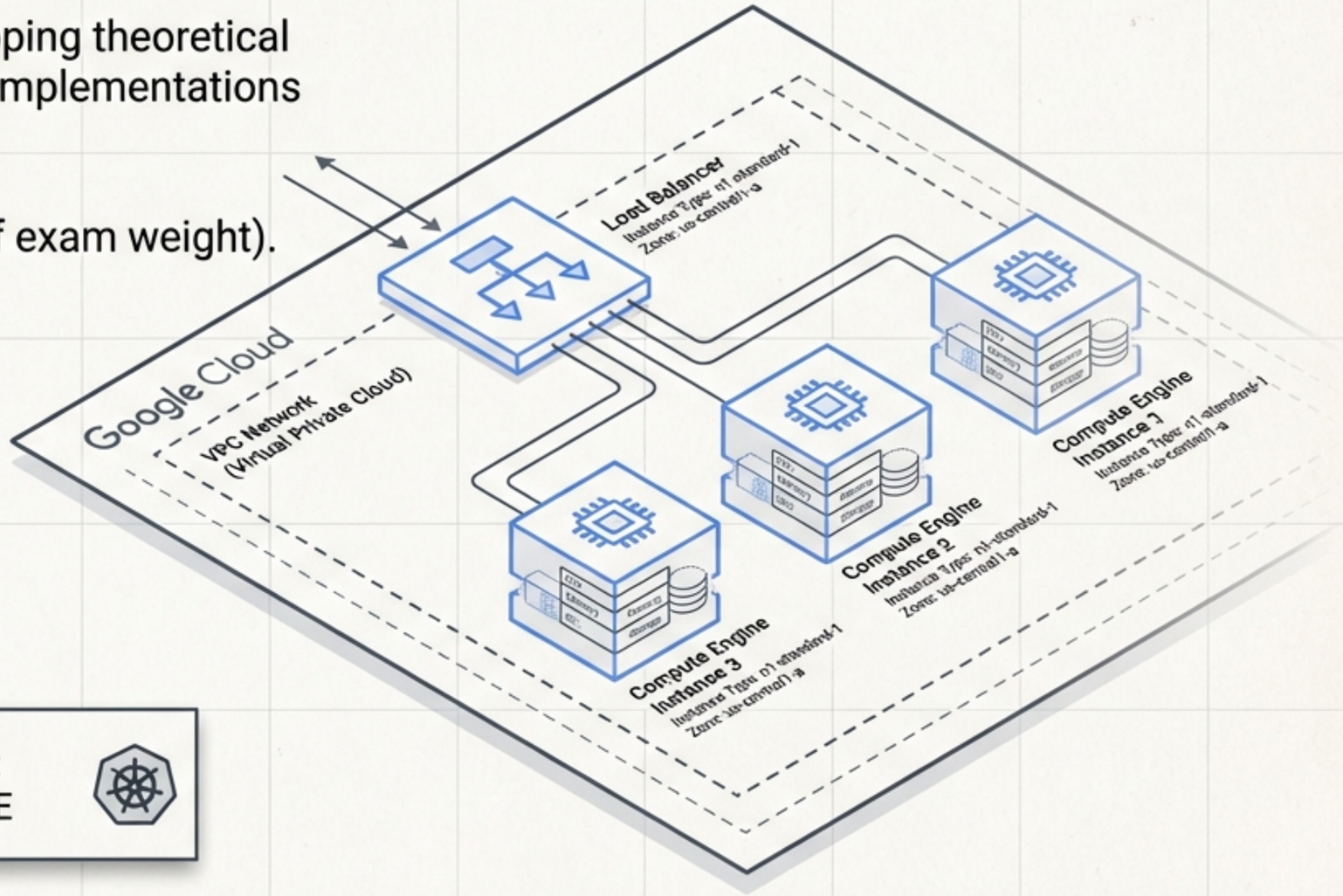
The Architect’s Blueprint for Infrastructure Provisioning

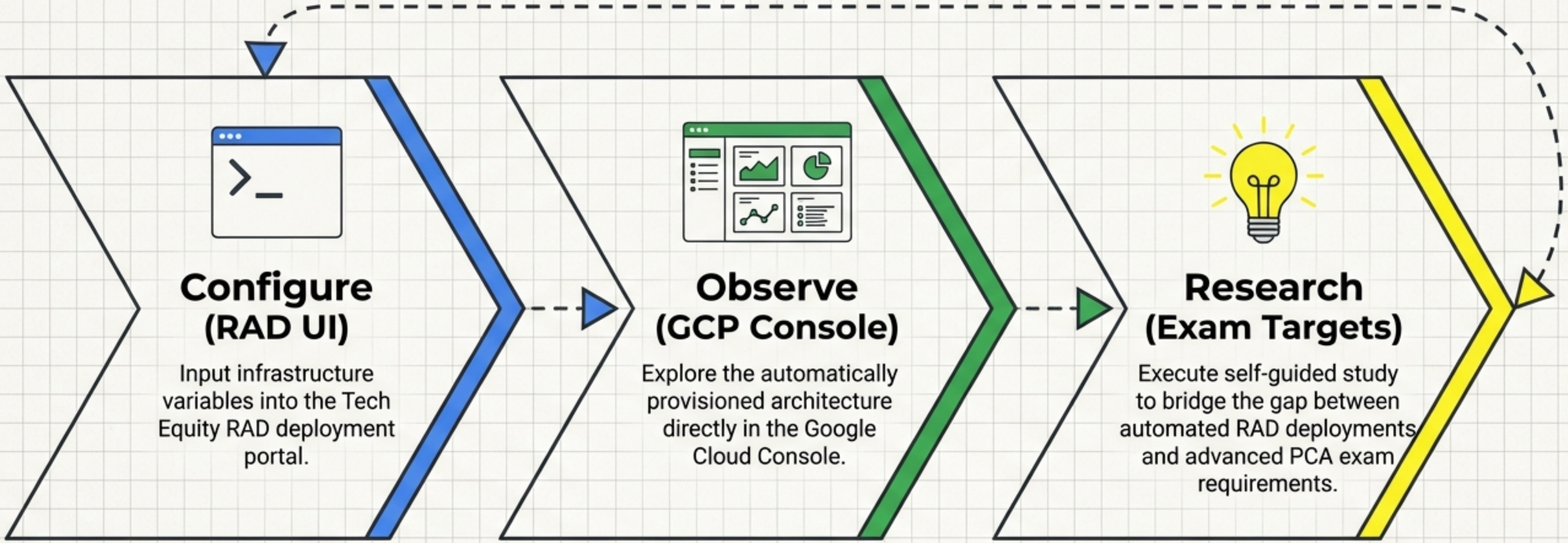
A highly structured visual reference mapping theoretical exam objectives to practical, real-world implementations using the Tech Equity RAD platform.

Targeting PCA Exam Section 2 (17.5% of exam weight).



- Module: GCP Services
- Module: App CloudRun
- Module: App GKE

The Action-Observation-Study Methodology



Visual Legend

- RAD Implementation
- 💡 Exam Target / Self-Study Sidebar

RAD Implementation

Establishing Custom-Mode VPCs and the Application Edge

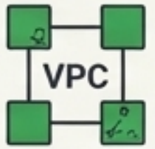
RAD UI Inputs

```
Group 2 Configuration:

availability_regions = [europe-west1, us-central1]
subnet_cidr_range = 10.0.0.0/16
```

Console Outputs

VPC network > VPC networks

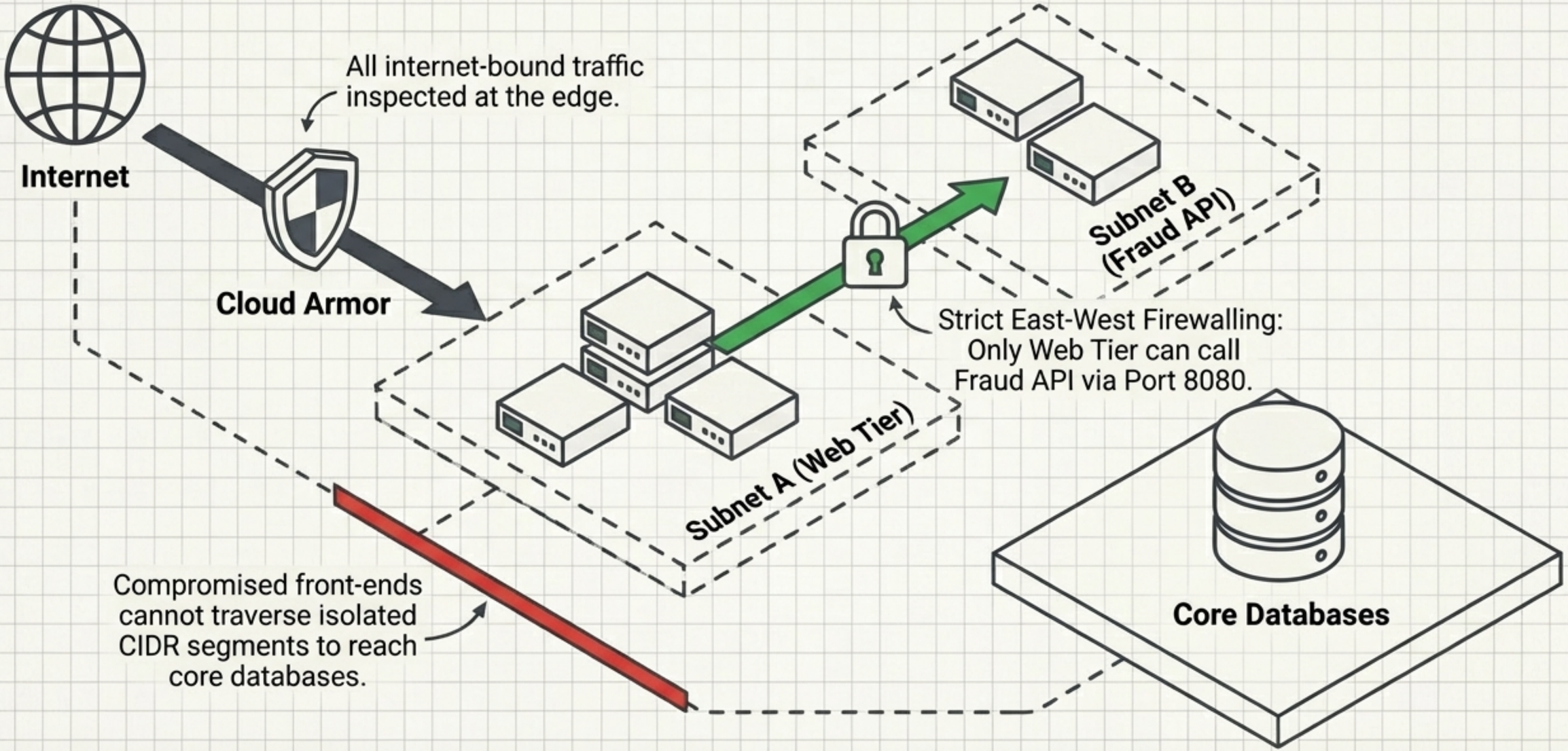


Custom-mode VPC creation enforcing complete control over IP spaces to prevent overlapping CIDR blocks.

Network Security > Cloud Armor

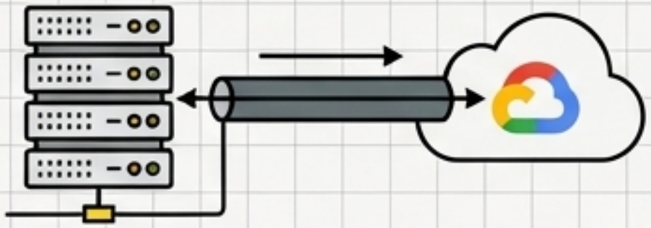
WAF rules providing immediate access control and firewalling at the edge.

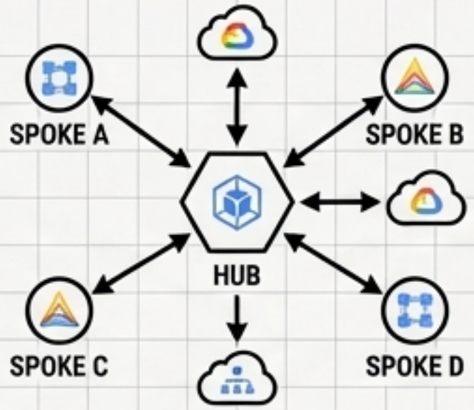
Defense-in-Depth for Financial Microservices

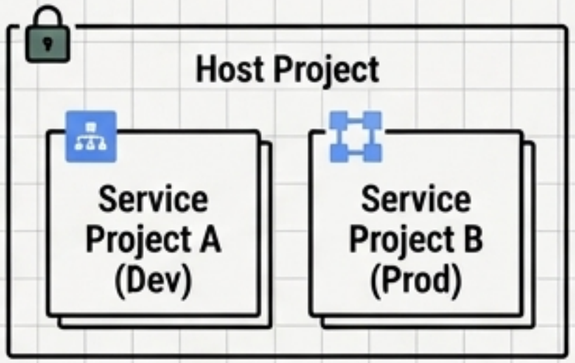


Exam Target

Networking Topologies Beyond the RAD Environment







Hybrid Networking

Cloud VPN (HA VPN) vs. Cloud Interconnect (Dedicated/Partner). Master BGP routing fundamentals to securely extend on-premises environments.

Multicloud Communication

Network Connectivity Center (NCC) for hub routing. Cross-Cloud Interconnect for physical multi-cloud links. GKE Enterprise (Anthos) for cross-cloud Kubernetes control planes.

Shared VPCs

Centralize network administration by logically dividing resources and billing between overarching Host projects and isolated Service projects.

RAD Implementation

Declarative Storage Allocation and Automated Protection

UI INPUT VARIABLES

`storage_buckets (Group 10 & 17)`

`Automated Workload Identities`

`backup_schedule (Group 6)`

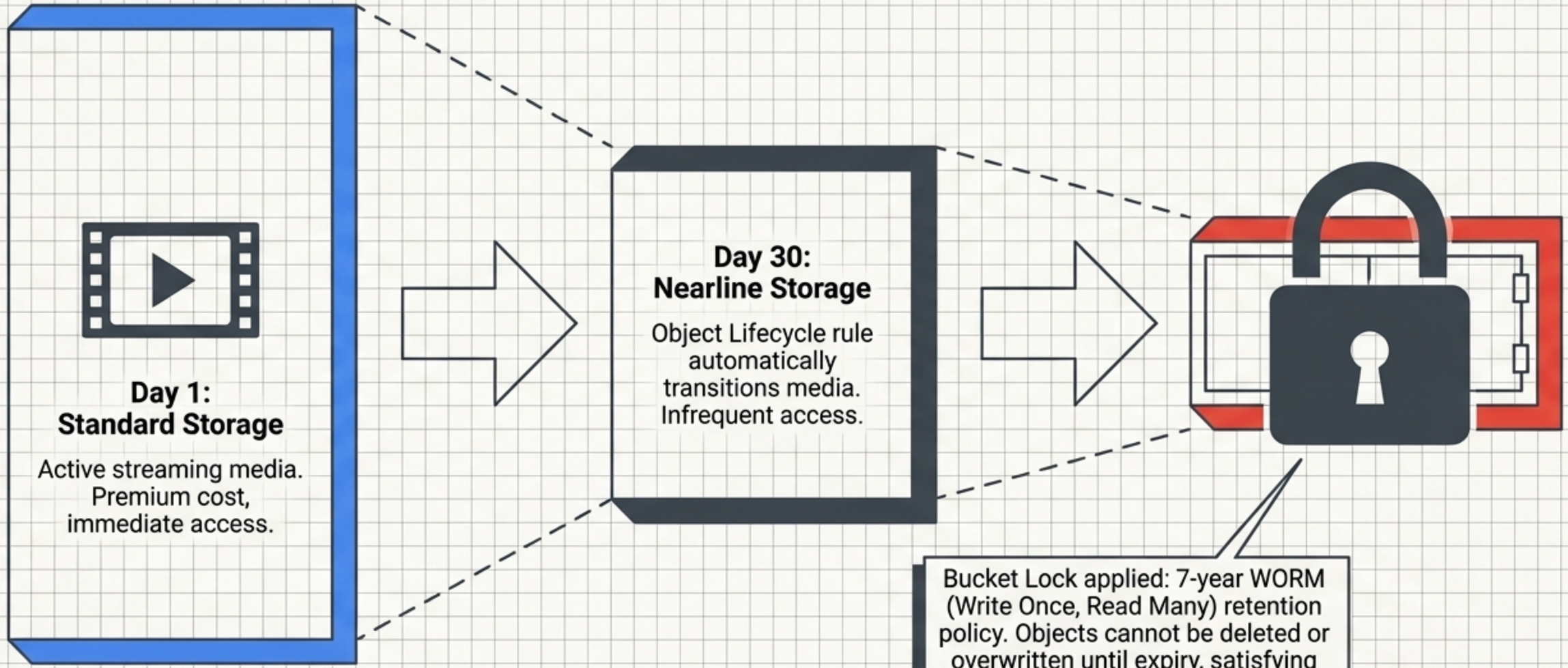
GCP OUTCOMES

Cloud Storage > Buckets
Verifying regional/multi-regional storage classes.

IAM & Admin > Service Accounts
Assigning minimum privileges like roles/storage.objectAdmin.

Automated relational database backups pushed securely to Cloud Storage.

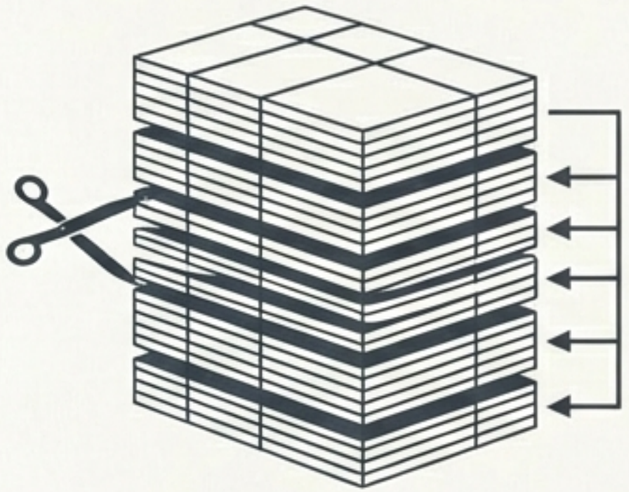
Media Archive Lifecycle and SEC/FINRA Compliance



Exam Target

Petabyte-Scale Data Transfer and Growth Strategies

	Online / API Driven
Online / API Driven	**Storage Transfer Service (STS)** Ideal for network-based migrations from other clouds or on-prem with high bandwidth.
Offline / Hardware	**Transfer Appliance** Physical hardware shipped to your data center. Ideal for massive datasets with constrained network bandwidth.



BigQuery Data Growth

Plan for petabyte-scale growth using BigQuery data warehousing. Master partitioning and clustering strategies to optimize query performance and control costs.

RAD Implementation

Abstracting Infrastructure via Declarative Provisioning

deploy_application
(Group 3)

Serverless Computing
(Cloud Run)

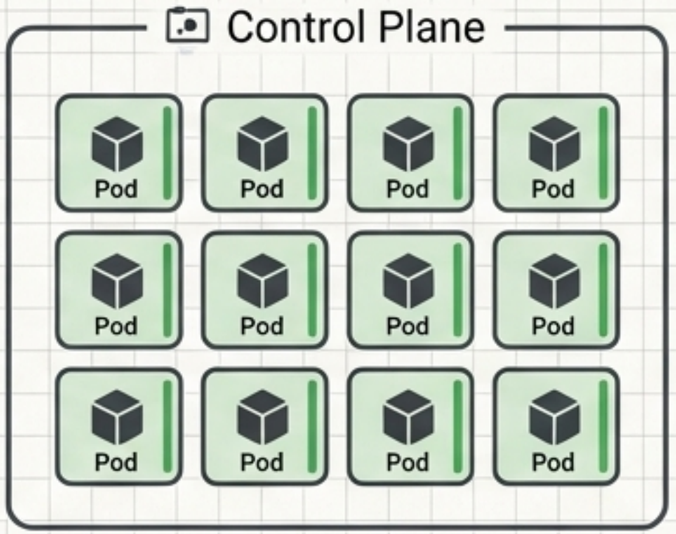
Driven by:
`container_image` & `container_port`

Scaled by: `min_instance_count`

Container Orchestration
(GKE Autopilot)

Driven by: `container_resources`
(Pod requests)

Underlying node infrastructure is fully abstracted by the Autopilot cluster.



The Compute Workload Decision Matrix

Cloud Run	GKE Autopilot	Spot VMs (💡 Exam Target)
Profile Serverless, variable burst traffic.	**Profile** Orchestrated, stateful data-processing pipelines.	**Profile** 91% cost savings vs standard VMs. No fixed 24-hour maximum lifespan.
Configuration `min_instance_count: 0` for dev cost savings; `min_instance_count: 2` for production to eliminate cold-start latency.	**Configuration** Requires persistent volumes and fine-grained CPU/GPU resource control.	**Use Case** Batch processing, stateless rendering, fault-tolerant workloads that survive sudden reclamation.

💡 Additional Compute Targets

1. **OS Patch Management**: Automating fleet-wide VM patching.
2. **VMware Engine**: Native vSphere lift-and-shift without refactoring.

Exam Target

Orchestrating End-to-End ML Workflows

Kubeflow / TensorFlow Extended (TFX) Orchestration

Data Ingestion

Pulling structured training datasets natively from BigQuery.

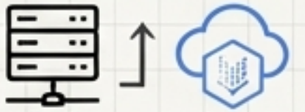
Model Training

Executing automated training runs.

Evaluation

Validating model accuracy and performance against baselines.

Deployment



Pushing to managed endpoints for consumption.

Context:

The RAD modules deploy standard web applications. PCA candidates must independently master automating this ML pipeline using Vertex AI Pipelines.

Exam Target

Machine Learning Serving and Infrastructure Architecture

Serving Types

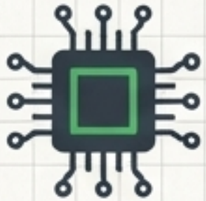
Online Serving (Vertex AI Feature Store)

- Low latency, per-entity lookups.
- Prevents training-serving skew.

Batch Serving (BigQuery ML)

- High throughput, full dataset predictions.

Compute Types



AI Hypercomputer (GPUs/TPUs)

- Integrating GPUs/TPUs within Vertex AI for large-scale training.
- Optimize via on-demand vs. reserved capacity.

Serverless Inference (Cloud Run)

- Using Cloud Run functions as lightweight inference endpoints for lower-volume, pre-trained models.

💡 Exam Target

The Six Pillars of Prebuilt Google AI APIs

Vision

Cloud Vision API: Image classification, OCR, object detection, face detection.

Image

Imagen API: Image generation and editing from text prompts.

Video

Video Intelligence API: Shot detection, label detection, object tracking, transcript extraction.

Audio

Speech-to-Text (transcription) & Text-to-Speech (synthesis).

Conversation

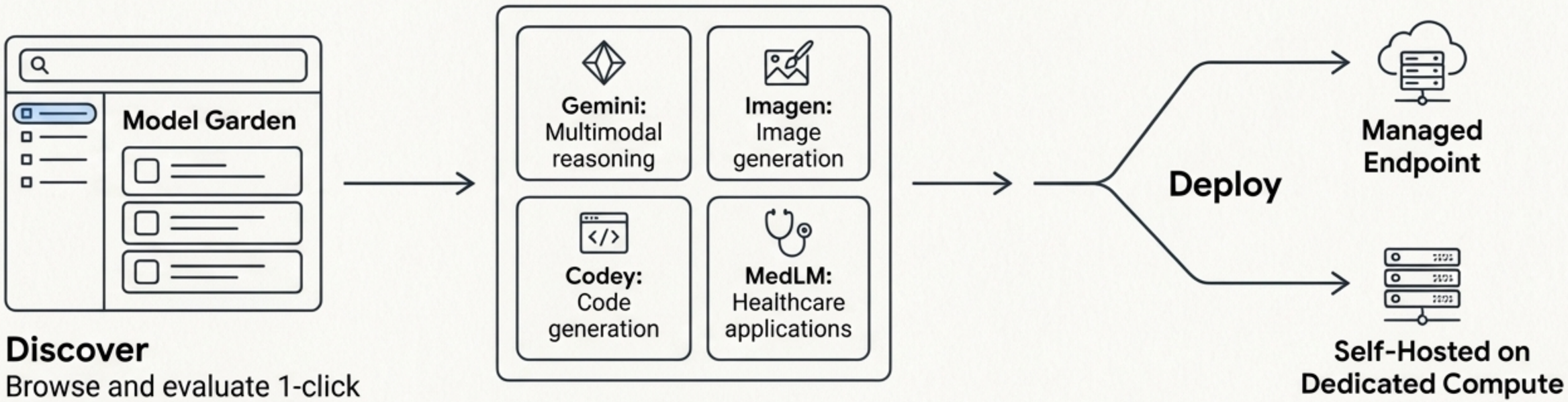
Dialogflow CX: Enterprise-grade virtual agents and complex conversational AI flows.

Search

Vertex AI Search: Enterprise search spanning structured and unstructured data sources.

💡 Exam Target

Model Garden and Enterprise AI Deployment



Discover
Browse and evaluate 1-click foundation models in the Model Garden console.

Stage 2: Select
Select Google-Native Models

Stage 3: Deploy

Enterprise Integration: Enhance enterprise workflows by integrating AI Agents and NotebookLM into the overarching architecture without building from scratch.

Synthesis: Section 2 Exam Readiness Checklist

	RAD Mastered (Configured via UI)	Self-Study Required (Exam Target)
2.1 Networking	Custom VPCs, Edge Security, Subnet Routing.	Hybrid VPN/Interconnect, NCC Multicloud, Shared VPCs.
2.2 Storage	Bucket Allocation, IAM Object Security, Automated Backups.	STS vs Transfer Appliance, WORM Compliance, BigQuery Partitioning.
2.3 Compute	Cloud Run serverless scaling, GKE Autopilot orchestration.	Spot VM lifecycles, OS Patch Management, VMware Engine.
2.4 ML Workflows	N/A (Web App Focus).	Kubeflow/TFX pipelines, Feature Store vs BigQuery ML serving, GPU/TPU tuning.
2.5 AI APIs	N/A.	The 6 API Pillars, Model Garden deployments, Gemini enterprise integration.

Deploy the Tech Equity RAD platform today and begin your targeted research.