

25–25%

Designing a cloud solution infrastructure that meets business requirements

- La estrategia de producto parte del valor de negocio: el coste total de propiedad incluye la gestión, por lo que conviene priorizar servicios gestionados y sin servidor como Cloud Run. Para casos de uso concretos hay productos
- listos (Quisitos en Amazon AWS (disponibilidad, resiliencia, elasticidad) de la misma manera que los de estas alternativas) y otras regiones, de tipo regional en los casos de alta disponibilidad se definen en métricas de experiencia de usuario,
- como la operación de peticiones por segundo y de latencia (tiempo de respuesta) y la estabilidad de los datos (impacto de diseño el RPO (tiempo máximo sin servicio) y el RPO (periodo máximo de pérdida de datos)). Los patrones en frío, templado
- y la optimización de costes es una prioridad de cualquier negocio y se apoyan en la visibilidad (etiquetas, exportación de facturación a BigQuery) y en ajustar los recursos a las necesidades reales y en los descuentos adecuados: compromisos de
- reserva y pagar los diseños de las aplicaciones conviene de las opciones de servicios como Amazon AWS (EBS, EBS, EBS) de
- vitales Gestión de datos como MemoryStore para lecturas rápidas, diseñar para la degradación gradual (limitar la frecuencia y descartar el exceso en el frontend) y seguir el recorrido de las peticiones con Cloud Trace para localizar

Designing a cloud solution infrastructure that meets technical requirements

- El **Veeva Architect** Framework agrupa sus recomendaciones en seis pilares (excelencia operativa, seguridad, fiabilidad, coste, rendimiento y sostenibilidad) y en perspectivas transversales como IA y ML. Entre sus principios:
- la **alta disponibilidad se diseña replanteando el 100% de cada crítico** en varios dominios de carga, con **consultas automatizadas de SQL** en configuración y **responsabilidad del cliente sobre la eficiencia** de los recursos.
- Google Cloud permite **adaptar el High Availability para Cloud RDS** escalando y colocando de forma automática tipos de **Cláusulas Personalizadas** de **RDS** dentro de regiones, licencias por CPU, Hyperdisk cambia IOPS y
- Pandicore convierte **el almacenamiento horizontal y las operaciones de estado** en **grupos de instancias de alta disponibilidad** con **grupos de datos predictivos** basados en **punto de control** y **carga**, escalador automático de clústeres de GKE
- **basado en la resolución de problemas** y **encontrando los usuarios** Cloud VPN requiere **padre de aplicación** base regional, redes escalables regionales, **charPhy** y **fast**, el nivel Premium para entrar en la red de Google por el punto de presencia más cercano. Dentro de una zona, la colocación compacta reduce la latencia entre VMs

Designing network, storage, and compute resources

- Paralelismo entre los canales de entrada para los datos temporales en Cloud VPN (menor coste y ancho de banda) y Cloud Interconnect: Dedicated si hay presencia en una instalación de colocation, Partner a través de un proveedor
- Genoss Entrepreneur Agenda Platform HVA / PNA los modelos Coremia Model Garden (modelo de Google, los socios viajan en la política de Evange Gizaglé Apasá para los personales) la plataforma con RAG para reducir alucinaciones y
- Agave VR Compartida Sessaliza, Superdes Banks para el género se en el proyecto Al happy y delegar el uso de subredes, software abierto computación de los usuarios por el mismo WolkRAch es el dato para el aprendizaje de los gráficos;
- Právile Sejeicé Canalizpolinres de Apache Beam posulines y costil Painting de Jopias dals poligestiona cada fegos jerárquicos del poder que las combinarse proye Pub/Sol para la incingesta Y BigQ NAT de salida simo Managed Service for
- Apache Spark frente a la arquitectura en la bodega (B-Sistema Distribuido) para el sistema SSD de archivo Data Filestore, NeptApoclonlines de la base de datos y para trabajos EgoQue Storage de la Base de datos, Bigtable sirve datos con

Creating a migration plan (i.e., documents and architectural diagrams)

- [illegible]

financiero, el coste total de propiedad debe in

Envisioning future solution improvements

- **Elasignifica el poderse comenzar a optimizar flexibles** el ciclo continuo de iteraciones con objetivos medibles, como SLOs, y evaluar las habilidades del equipo antes de adoptar servicios nuevos. Herramientas como las recomendaciones de **ActiveAssist** de este lectora cambian, así que la reconfiguración de la **AntidBilling** a **BigQuery** en capacidad antes de tenerlas.
- **Envejecer y nuevos** el **AntidBilling** de **BigQuery** en capacidad antes de tenerlas. Envejecer y nuevos el **AntidBilling** de **BigQuery** en capacidad antes de tenerlas.
- **Y reconfiguración de la** **AntidBilling** de **BigQuery** en capacidad antes de tenerlas. Y reconfiguración de la **AntidBilling** de **BigQuery** en capacidad antes de tenerlas.

2. Managing and provisioning a cloud solution infrastructure

17-18%

Configuring network topologies

- Cloud VPN cifra el tráfico por internet; Dedicated Interconnect exige equipos en una instalación de colocation y Partner Interconnect conecta mediante un proveedor. Cloud Interconnect no cifra por defecto: HA VPN sobre Cloud Interconnect añade IPsec. Para DNS híbrido, las zonas de reenvío consultan servidores on-premises y la política de servidor con reenvío entrante permite consultar zonas privadas desde on-premises.
- Cross-Cloud Interconnect aprovisiona una conexión física dedicada entre Google y otro proveedor de nube, con un par de conexiones por proveedor. Entre redes de Google Cloud, el emparejamiento de VPC no es transitivo; Google recomienda radios de VPC de Network Connectivity Center, incluso entre organizaciones, y Private Service Connect.
- Las políticas de cortafuegos jerárquicas se asocian a la organización o a carpetas, no se pueden anular desde niveles inferiores y delegan con goto_next. Cloud NGFW Enterprise inspecciona el tráfico de VMs y GKE con endpoints de cortafuegos y perfiles de prevención de amenazas. Cloud Armor aplica reglas WAF en el borde de los balanceadores externos, e IAP permite SSH a VMS sin IP externa según la identidad.
- La VPC compartida centraliza subredes y cortafuegos en un proyecto host; el rol de usuario de red se concede por subred. Los balanceadores regionales terminan TLS solo en su región y los globales en ubicaciones de todo el mundo. Cloud NAT da salida a internet con IPs fijas sin aceptar conexiones entrantes, y el acceso privado a Google

Configuring individual storage systems

- Google Cloud ofrece almacenamiento en bloques (Persistent Disk, Hyperdisk, Local SSD efímero), de archivos (Filestore para NFS compartido) y de objetos (Cloud Storage). Hyperdisk permite fijar y cambiar IOPS o rendimiento con el volumen en uso; Hyperdisk ML se conecta en solo lectura a muchas instancias, y los Storage Pools comparten capacidad y rendimiento para evitar el sobreaprovisionamiento.
- Dataflow ejecuta canalizaciones Apache Beam por lotes y streaming sin gestionar VMs, con autoescalado y procesamiento exactamente una vez. Managed Service for Apache Spark (antes Dataproc) ofrece clústeres o ejecución serverless que solo cobra mientras corre la carga. Con los datos en Cloud Storage, los trabajadores secundarios (también spot) escalan solo la computación y los clústeres inactivos se eliminan y recrean.
- El acceso uniforme a nivel de bucket desactiva las ACLs y deja solo IAM; la prevención de acceso público, aplicable con la restricción storage.publicAccessPrevention, impide hacer públicos los datos. Las URLs firmadas dan acceso temporal a un objeto sin cuenta de Google. En BigQuery, las etiquetas de política restringen columnas y las políticas de Transfer Appliance encaja con grandes volúmenes y poco ancho de banda; Storage Transfer Service mueve datos desde otras nubes o desde NFS on-premises a gran escala con límite de ancho de banda, y gcloud storage sirve para transferencias puntuales. Cloud CDN sirve contenido desde el borde de Google, y los buckets birregionales.
- Object Lifecycle Management cambia la clase de almacenamiento o borra objetos según condiciones como la antigüedad. Una política de retención impide borrar objetos antes de tiempo y Bucket Lock la hace permanente; las retenciones basadas en eventos reinician el periodo al liberarse. En BigQuery, la caducidad de partición elimina

Configuring compute systems

- Las plantillas de instancia definen la configuración de las VMs y son obligatorias para los MIG; las regionales aíslan errores y ayudan a la residencia de datos. Las reservas aseguran capacidad: las de consumo automático las usa cualquier VM compatible y las de destino específico solo las VMs que apuntan a ellas. La creación masiva elige zonas con capacidad y valida la cuota, pero no crea un grupo gestionado.
- Las Spot VMs usan capacidad sobrante con grandes descuentos, pero Compute Engine puede expulsarlas en cualquier momento; al expulsarlas las detiene o, con la acción DELETE, las elimina. Un aviso de expulsión de 120 segundos da margen a cargas que necesitan más de 30 segundos para cerrar. En GKE conviene aplicar taints a los nodos Spot, usar nodeSelector y tolerancias, y conservar un grupo de nodos estándar para el sistema.
- Cloud Run llega a la VPC con la salida directa de VPC (recomendada, sin cargos de VM y con etiquetas de red por servicio) o con conectores; para una IP de salida estática, todo el tráfico debe pasar por Cloud NAT. Los clústeres de GKE nativos de VPC usan rangos secundarios para Pods y Services, ampliables con CIDR multi-Pod discontinuo.
- Infrastructure Manager ejecuta Terraform de forma gestionada desde Git o Cloud Storage, guarda el estado de VMware Engine se conecta a la VPC mediante emparejamiento con su red de VMware Engine.
- Y, mediante políticas de SO, mantiene el software deseado en grupos de VMs filtrados por etiquetas. Los MIG despliegan nuevas plantillas con actualizaciones graduales controladas por maxSurge y maxUnavailable.
- Los clústeres regionales de GKE replican el plano de control en varias zonas y mantienen la API disponible durante las actualizaciones; Autopilot siempre es regional. Las ventanas y exclusiones de mantenimiento controlan cuándo y qué se actualiza, por ejemplo Sin actualizaciones secundarias. El escalador automático de clústeres decide según

Leveraging Gemini Enterprise Agent Platform for end-to-end ML workflows

- Agent Platform Pipelines orquesta sin servidor flujos de ML definidos como DAG con Kubeflow Pipelines o TFX, compilados a YAML y ejecutables puntualmente o con la API del programador. ML Metadata guarda el linaje de artefactos de cada ejecución. La caché de ejecución omite pasos con la misma clave y se desactiva por tarea o por trabajo.
- Agent Platform Feature Store usa tablas o vistas de BigQuery como almacén offline y sirve los valores más recientes online mediante vistas de características; registrar grupos con marca de tiempo permite servir solo el último valor.
- AutoML necesita un conjunto de datos gestionado, al que Agent Platform accede con un agente de servicio. Antes de entrenar hay que evitar la fuga de datos y el sesgo entre entrenamiento y servicio.
- Al Hypercomputer integra GPUs y TPUs, software abierto y modelos de consumo: flex-start con Dynamic Workload Scheduler para trabajos que pueden esperar, Spot para cargas tolerantes a fallos y reservas para máxima garantía de capacidad. Las TPUs encajan con modelos matriciales grandes y se escalan en slices. Cloud Run ofrece GPUs

Configuring prebuilt solutions or APIs with Agent Platform

- Cada API de IA de Google cubre un tipo de contenido: Vision analiza imágenes (etiquetas, logotipos, OCR), Video Intelligence analiza vídeo (planos, logotipos, objetos), Speech-to-Text transcribe audio de forma síncrona, asíncrona o en streaming y Text-to-Speech sintetiza voz. Document AI extrae campos y tablas de documentos. Natural Language analiza texto. Cloud Translation traduce y Agent Search ofrece búsqueda y respuestas generativas.
- Gemini Enterprise une búsqueda en la intranet, asistente y agentes sobre datos conectados mediante almacenes de datos y conectores prediseñados (Jira, Confluence, SharePoint), con respuestas que respetan los permisos y acciones para crear o actualizar registros. Gemini Notebook Enterprise, antes NotebookLM Enterprise, resume e investiga fuentes subidas como copias estáticas, con los datos en el proyecto y uso compartido solo dentro de él.
- Model Garden reúne modelos de Google, de partners y abiertos con un patrón de despliegue común: los modelos gestionados (MaaS) son sin servidor, mientras que los autodesplegados se ejecutan en tu proyecto y tu VPC y pueden usar reservas de Compute Engine. La política vertexai.allowedModels limita qué modelos y acciones se permiten, y los endpoints regionales o multirregionales ayudan a respetar la residencia de datos.

3. Designing for security and compliance

17-18%

Designing for security

- IAM combina políticas de permisos que se heredan por la jerarquía, políticas de denegación que se evalúan antes y políticas de Principal Access Boundary que limitan a qué recursos pueden acceder las entidades. Las buenas prácticas son evitar roles básicos en producción, conceder roles predefinidos en el ámbito más pequeño, usar una cuenta de servicio por componente y condiciones para que el acceso caduque solo.
- La jerarquía de recursos va de la organización (raíz, ligada a Cloud Identity o Google Workspace) a carpetas opcionales y proyectos. Las políticas de permisos, de denegación y de organización se heredan hacia abajo; los permisos de niveles inferiores son aditivos y solo una política de denegación los restringe. Google recomienda un único nodo raíz y elegir la estructura por entornos, filiales o responsabilidad, no por organigrama.
- Google Cloud cifra por defecto los datos en reposo y en tránsito; para más control, Cloud KMS gestiona claves cuyo material no se puede exportar, Cloud HSM las aloja en HSM gestionados por Google y Cloud EKM las mantiene en un gestor externo. Secret Manager guarda contraseñas y claves de API con versiones, IAM por secreto y rotación programada mediante avisos de Pub/Sub.
- La separación de funciones evita que una sola entidad tenga todos los permisos para una acción maliciosa. En Cloud KMS separa a los gestores de claves de sus usuarios (idealmente solo el agente de servicio), con proyectos de claves dedicados o políticas de denegación; Privileged Access Manager añade acceso temporal con aprobación de otra persona, y el superadministrador de Workspace delega Google Cloud en Organization Administrator.
- Los controles de seguridad se combinan en capas: VPC Service Controls crea perímetros que frenan la exfiltración desde las API de Google, las políticas de cortafuegos jerárquicas imponen reglas a toda la organización o a carpetas que los proyectos no pueden anular, las políticas de organización limitan configuraciones como ubicaciones o IP externas, y los registros de acceso a datos, desactivados por defecto salvo en BigQuery, deben habilitarse.

Designing for compliance

- El cumplimiento normativo es una responsabilidad compartida: tú eres dueño de tus datos y de muchos controles de privacidad. Para HIPAA, Google firma un Business Associate Agreement que cubre toda su infraestructura; Assured Workloads aplica paquetes de controles de residencia y de acceso del personal a carpetas; y la soberanía se refuerza con claves externas (Cloud EKM), Key Access Justifications y Access Transparency.
- Para datos de tarjetas, el alcance PCI DSS incluye todo sistema que almacene, procese o transmita datos de titulares y los sistemas conectados; se reduce con carpetas para usuarios designados, una VPC dedicada en modo personalizado con VPC Service Controls y tokenización. Para PII, Sensitive Data Protection descubre datos con perfiles a nivel de organización y los desidentifica con enmascaramiento, hash o tokenización reversible.
- Google Cloud solo emite informes SOC 2 Tipo II, elaborados por un auditor independiente sobre sus propios controles, y se descargan bajo demanda en Compliance Reports Manager; las cartas puente cubren el hueco hasta el cierre del periodo del cliente. Por la responsabilidad compartida, tus controles los evidencias tú: comprueba el alcance de cada servicio y supervisa tu configuración con el marco SOC 2 de Compliance Manager.
- Cloud Audit Logs registra actividad de administración y eventos del sistema siempre; los registros de acceso a datos están desactivados salvo en BigQuery y hay que habilitarlos. Para auditoría centralizada, usa un receptor agregado de organización con --include-children hacia un bucket de registros bloqueado con retención. Access Transparency registra los accesos del personal de Google.

4. Analyzing and optimizing technical and business processes

15-15%

Analyzing and defining technical processes

- Google Cloud protege todo el ciclo de vida del software: Cloud Workstations da entornos de desarrollo gestionados y coherentes; Artifact Registry centraliza dependencias, con repositorios virtuales que priorizan los paquetes privados; Assured OSS ofrece paquetes verificados; Artifact Analysis genera SBOM y detecta vulnerabilidades; Cloud Build produce procedencia no falsificable alineada con SLSA.
- Cloud Build es el servicio gestionado de CI: sus activadores compilan y prueban ante envíos o solicitudes de extracción. Cloud Deploy gestiona la entrega continua a Cloud Run y GKE con una secuencia de promoción, aprobaciones, estrategias canary automatizadas y reversión a la última versión correcta. Binary Authorization bloquea en el despliegue las imágenes sin las atestaciones exigidas.
- Para el análisis de causa raíz, elige la herramienta según el síntoma: Cloud Trace para la latencia entre servicios, Cloud Profiler para el consumo de CPU y memoria por función en producción y Connectivity Tests para fallos de red tras cambios de configuración. Después, haz una revisión posterior al incidente sin culpables, con cronología, causa inmediata y factores sistémicos, y compártela.
- Prueba barato primero: emuladores locales de Pub/Sub o Spanner para pruebas funcionales, sabiendo que el de Spanner no admite IAM ni refleja el rendimiento real. En Terraform, análisis estático con terraform validate, pruebas de módulos aisladas en proyectos nuevos y limpieza con terraform destroy. La recuperación se valida simulando fallos con carga en un entorno de preproducción y midiendo frente al RTO.
- Service Catalog permite a los administradores publicar soluciones seleccionadas (Terraform, plantillas o enlaces) en catálogos compartidos con organizaciones, carpetas o proyectos; los usuarios con Catalog Consumer las lanzan, se respetan las políticas de la organización y actualizar una solución no cambia los despliegues existentes. Infrastructure Manager ejecuta Terraform gestionado con revisiones y vistas previas.

Analyzing and defining business processes

- Implica a todas las partes interesadas desde el inicio para acordar alcance y decisiones, y mantenlas informadas con objetivos, puntos de control e hitos. Para influir en finanzas, argumenta con el coste total de propiedad y el valor de negocio, no con el precio de catálogo. Elige equipos pioneros motivados y patrocinados por la dirección, y recoge opiniones con encuestas y entrevistas.
- La gestión de cambios combina gobierno (aprobaciones y comunicación), evaluación de riesgos, pruebas y despliegues controlados con reversión. En Google Cloud se apoya en Terraform con control de versiones y CI/CD, aprobaciones de Cloud Deploy integradas por Pub/Sub con herramientas de ITSM, políticas de organización probadas fuera de producción y aplicadas poco a poco, y cambios pequeños y frecuentes medidos con DORA.
- La preparación del equipo se evalúa con el Google Cloud Adoption Framework: cuatro temas (Learn, Lead, Scale y Secure) en fases táctica, estratégica o transformacional. La evaluación incluye formar por roles, dar entornos de prueba para experimentar y elegir equipos pioneros motivados, con prácticas DevOps y SRE y patrocinio. Realojar conserva las competencias actuales; refactorizar o reconstruir exige otras nuevas.
- Para decidir con criterio, Google recomienda registros de decisiones de arquitectura (ADR) que recojan opciones, requisitos y la decisión, guardados junto al código y con el historial de cambios. El plan de migración incluye un documento de diseño técnico, una matriz RACI y una cronología. Las pruebas de concepto con alcance definido e impacto de negocio medible aportan evidencias a la dirección.
- La relación con Cloud Customer Care exige elegir bien la prioridad: P1 si producción no se puede usar y no hay solución alternativa rápida, P2 si está degradada pero hay una. El escalado sirve para fallos del proceso, no para acelerar casos. Un buen caso incluye marcas ISO 8601, ID de proyecto y zona. Personalized Service Health avisa de incidencias que afectan a tus proyectos.

5. Managing implementation

12-13%

Advising development and operation teams to ensure the successful deployment of the solution

- En los grupos de instancias gestionados, maxSurge y maxUnavailable controlan el ritmo de una actualización proactiva; con maxUnavailable en 0 y maxSurge mayor que 0 no se pierde capacidad, y --canary-version prueba una plantilla en un subconjunto. Cloud Deploy promociona versiones por destinos, admite aprobaciones con requireApproval, despliegues canary en GKE y Cloud Run y reversiones a la última versión correcta.
- Apigee es la plataforma de gestión de API de Google Cloud: un proxy entre clientes y backends que aplica políticas de seguridad, tráfico y caché. Quota limita llamadas por aplicación o desarrollador en periodos largos con recuento preciso; SpikeArrest protege el backend frente a picos. Los productos de API agrupan operaciones y cuotas, y las claves se emiten para ellos. La lógica compleja va fuera de la pasarela.
- Los emuladores (Pub/Sub, Spanner) permiten pruebas unitarias y de integración locales, pero no admiten IAM y su rendimiento no es comparable con producción. Cloud Build ejecuta las pruebas en una plataforma coherente y bajo demanda en cada commit, y Locust o JMeter cubren el rendimiento. Para las pruebas de carga en GKE, mide primero la capacidad de un Pod sin autoescalado y genera la carga desde la misma región.
- Migration Center descubre recursos, calcula el TCO y detecta dependencias. Migrate to Virtual Machines replica VMs en segundo plano, permite test-clones en un entorno aislado y hace el cut-over con una breve parada. Database Migration Service migra a Cloud SQL y AlloyDB con replicación continua y conversión de esquemas. Storage Transfer Service mueve datos en línea; Transfer Appliance, cuando falta ancho de banda.
- Gemini Cloud Assist es el colaborador con IA de Google Cloud para diseñar, desplegar, diagnosticar y optimizar. Genera comandos de gcloud, manifiestos y Terraform, ejecuta investigaciones sobre registros y métricas, se activa desde alertas, traspasa casos al soporte y explica anomalías de costes. En segundo plano actúa con una cuenta de servicio dedicada auditada, y sus resultados deben validarse antes de usarse.

Interacting with Google Cloud programmatically

- Cloud Shell ofrece una VM temporal gestionada por Google con la CLI de gcloud y otras herramientas; solo persiste el directorio \$HOME, que no se puede ampliar (para más espacio, Cloud Workstations). El modo efímero arranca más rápido sin persistencia. El Cloud Shell Editor, basado en Code OSS, integra Cloud Code, que da soporte en VS Code, IntelliJ y Cloud Shell al desarrollo para Kubernetes y Cloud Run.
- Google recomienda los comandos gcloud storage frente a gsutil, herramienta heredada sin las funciones nuevas de Cloud Storage. En automatización, los scripts de gcloud deben basarse en el código de salida y en --format y --filter, no en la salida predeterminada ni en stderr, y las configuraciones con nombre separan cuentas y proyectos. bq ofrece --dry_run para estimar bytes y --maximum_bytes_billed para limitar el coste.
- gcloud CLI ofrece emuladores locales en memoria para Bigtable, Spanner, Pub/Sub y Firestore; las bibliotecas de cliente los usan mediante variables como PUBSUB_EMULATOR_HOST o SPANNER_EMULATOR_HOST. Los emuladores son solo para desarrollo y pruebas: omiten IAM y la autenticación, no igualan el rendimiento de producción y tienen carencias, como no admitir gcloud pubsub ni las API administrativas de Bigtable.
- Google recomienda guardar el estado de Terraform en remoto en un bucket de Cloud Storage con versiones, estructurar cada servicio en módulos y una configuración raíz por entorno con su propio backend, y ejecutar Terraform desde un pipeline automatizado que guarde el plan para aprobarlo antes de aplicarlo. Infrastructure Manager ejecuta Terraform como servicio gestionado con revisiones y previsualizaciones.
- Las aplicaciones deben autenticarse mediante Application Default Credentials: una cuenta de servicio asociada en Google Cloud, ADC de gcloud (opcionalmente suplantando una cuenta de servicio) en las estaciones de trabajo y Workload Identity Federation fuera de Google Cloud, evitando las claves de cuenta de servicio. Las claves de API llevan restricciones de API y de aplicación, y solo se reintentan errores transitorios con retroceso exponencial.

6. Ensuring solution and operations excellence

12-13%

Understanding the principles and recommendations of the operational excellence pillar of the Google Cloud Well-Architected Framework

- El pilar de excelencia operativa pide definir SLO SMART seguidos con SLI en Cloud Monitoring, gestionar los incidentes con runbooks, seguimiento centralizado, corrección automatizada y revisiones posteriores sin culpables, y ajustar y escalar los recursos con Cloud Monitoring y Recommender. La automatización, la IaC y el CI/CD reducen el trabajo repetitivo, y las retrospectivas impulsan la mejora continua.

Familiarity with Google Cloud Observability solutions

- Cloud Logging envía las entradas mediante receptores a buckets de registros, Cloud Storage, BigQuery o Pub/Sub; los receptores agregados de organización o carpeta centralizan los registros, y los receptores a Pub/Sub alimentan herramientas de terceros. El Ops Agent recoge registros, métricas y trazas de Compute Engine, y un ámbito de métricas multiproyecto concentra paneles y alertas de muchos proyectos.
- Cloud Profiler es un generador de perfiles estadístico y de baja sobrecarga que recoge de forma continua datos de CPU y heap en producción y los atribuye al código, mientras que Cloud Trace mide la latencia de extremo a extremo entre servicios e identifica dependencias lentas. El Well-Architected Framework recomienda pruebas de carga realistas y de estrés por encima del pico antes de los lanzamientos.
- Las políticas de alertas de Cloud Monitoring evalúan métricas, entradas de registro o resultados de consultas SQL y avisan mediante canales. Hay condiciones de umbral, de ausencia de métrica para series sin datos, de valor previsto para infracciones anticipadas y PromQL. Para los SLO, Google recomienda una alerta de consumo rápido con un periodo corto y otra de consumo lento con uno más largo.

Deployment and release management

- Cloud Deploy promociona versiones por destinos en una secuencia definida, con aprobaciones opcionales (requireApproval, anunciadas por Pub/Sub), reversiones que vuelven a desplegar la última versión correcta y estrategias canary en las que puede repartir el tráfico automáticamente. Cloud Run permite desplegar revisiones sin tráfico, probarlas con URL de etiqueta y migrar el tráfico de forma gradual por porcentaje.

Assisting with the support of deployed solutions

- Un buen caso de soporte da a Customer Care una marca de tiempo ISO 8601, el producto y el mecanismo de la solicitud, la zona, los ID de proyecto y artefactos, usa un caso por problema y sube a P1 el caso existente si afecta a producción. Personalized Service Health muestra y alerta de incidentes de Google Cloud relevantes para tus proyectos, y Error Reporting agrupa los errores de la aplicación.

Evaluating quality control measures

- El control de calidad en Google Cloud se automatiza en la canalización: Cloud Deploy ejecuta contenedores de verificación tras cada despliegue y un fallo detiene el rollout, y Binary Authorization solo admite imágenes con certificaciones de las etapas previas. Para el código de infraestructura conviene probar de menor a mayor coste: análisis estático, módulos aislados y, al final, extremo a extremo en un entorno nuevo.

Ensuring the reliability of solutions in production (e.g., chaos engineering, penetration testing, and load testing)

- La fiabilidad en producción se demuestra probando: simula fallos de forma gradual y con carga realista, comprueba que saltan las alertas y compara la recuperación con el RTO y el RPO, de forma periódica. Cloud Service Mesh inyecta retrasos y errores en un porcentaje de solicitudes; Web Security Scanner complementa, sin sustituirla, la revisión de seguridad manual; en Apigee, las pruebas de carga se avisan antes a soporte.