



Google Cloud

Mantenimiento y supervisión

Philipp Maier

Desarrollador de cursos, Google Cloud

En el último módulo del curso, abordaremos el mantenimiento y la supervisión de aplicaciones.

Objetivos de aprendizaje

- Administrar versiones nuevas de los servicios mediante actualizaciones progresivas, implementaciones azul-verde y lanzamientos Canary
- Prever, supervisar y optimizar el costo de los servicios con la calculadora de precios de Google Cloud, los informes de facturación y el análisis de datos de facturación
- Observar si sus servicios cumplen con los SLO utilizando Cloud Monitoring y los paneles
- Usar verificaciones de tiempo de actividad para determinar la disponibilidad de los servicios
- Responder ante las interrupciones del servicio mediante alertas de Cloud Monitoring

El mantenimiento aborda principalmente la forma en que se actualizan las aplicaciones en ejecución, las distintas estrategias disponibles y cómo las diferentes plataformas de implementación respaldan esas estrategias. Respecto a la supervisión, analizaremos este campo vital para las aplicaciones nativas de la nube desde dos puntos de vista:

1. En primer lugar, hablaremos sobre los costos para asegurarnos de que los recursos se aprovisionen según la demanda. En definitiva, ¿por qué tendría que pagar por recursos que no necesita?
2. En segundo lugar, veremos cómo utilizar Cloud Monitoring y paneles para implementar la supervisión y la observabilidad a fin de determinar el estado de los servicios y las aplicaciones, así como generar alertas al respecto.

Esto también nos permitirá definir verificaciones de tiempo de actividad y utilizar las alertas de Cloud Monitoring para identificar interrupciones del servicio. Comencemos.

Objetivos de aprendizaje

- Administrar versiones nuevas de los servicios mediante actualizaciones progresivas, implementaciones azul-verde y lanzamientos Canary
- Prever, supervisar y optimizar el costo de los servicios con la calculadora de precios de Google Cloud, los informes de facturación y el análisis de datos de facturación
- Observar si sus servicios cumplen con los SLO utilizando Cloud Monitoring y los paneles
- Usar verificaciones de tiempo de actividad para determinar la disponibilidad de los servicios
- Responder ante las interrupciones del servicio mediante alertas de Cloud Monitoring

El mantenimiento aborda principalmente la forma en que se actualizan las aplicaciones en ejecución, las distintas estrategias disponibles y cómo las diferentes plataformas de implementación respaldan esas estrategias.

Objetivos de aprendizaje

- Administrar versiones nuevas de los servicios mediante actualizaciones progresivas, implementaciones azul-verde y lanzamientos Canary
- Prever, supervisar y optimizar el costo de los servicios con la calculadora de precios de Google Cloud, los informes de facturación y el análisis de datos de facturación
- Observar si sus servicios cumplen con los SLO utilizando Cloud Monitoring y los paneles
- Usar verificaciones de tiempo de actividad para determinar la disponibilidad de los servicios
- Responder ante las interrupciones del servicio mediante alertas de Cloud Monitoring

Respecto a la supervisión, analizaremos este campo vital para las aplicaciones nativas de la nube desde dos puntos de vista:

En primer lugar, hablaremos sobre los costos para asegurarnos de que los recursos se aprovisionen según la demanda. En definitiva, ¿por qué tendría que pagar por recursos que no necesita?

Objetivos de aprendizaje

- Administrar versiones nuevas de los servicios mediante actualizaciones progresivas, implementaciones azul-verde y lanzamientos Canary
- Prever, supervisar y optimizar el costo de los servicios con la calculadora de precios de Google Cloud, los informes de facturación y el análisis de datos de facturación
- Observar si sus servicios cumplen con los SLO utilizando Cloud Monitoring y los paneles
- Usar verificaciones de tiempo de actividad para determinar la disponibilidad de los servicios
- Responder ante las interrupciones del servicio mediante alertas de Cloud Monitoring

En segundo lugar, veremos cómo utilizar Cloud Monitoring y paneles para implementar la supervisión y la observabilidad a fin de determinar el estado de los servicios y las aplicaciones, así como generar alertas al respecto.

Objetivos de aprendizaje

- Administrar versiones nuevas de los servicios mediante actualizaciones progresivas, implementaciones azul-verde y lanzamientos Canary
- Prever, supervisar y optimizar el costo de los servicios con la calculadora de precios de Google Cloud, los informes de facturación y el análisis de datos de facturación
- Observar si sus servicios cumplen con los SLO utilizando Cloud Monitoring y los paneles
- Usar verificaciones de tiempo de actividad para determinar la disponibilidad de los servicios
- Responder ante las interrupciones del servicio mediante alertas de Cloud Monitoring

Esto nos permitirá definir verificaciones de tiempo de actividad y

Objetivos de aprendizaje

- Administrar versiones nuevas de los servicios mediante actualizaciones progresivas, implementaciones azul-verde y lanzamientos Canary
- Prever, supervisar y optimizar el costo de los servicios con la calculadora de precios de Google Cloud, los informes de facturación y el análisis de datos de facturación
- Observar si sus servicios cumplen con los SLO utilizando Cloud Monitoring y los paneles
- Usar verificaciones de tiempo de actividad para determinar la disponibilidad de los servicios
- Responder ante las interrupciones del servicio mediante alertas de Cloud Monitoring

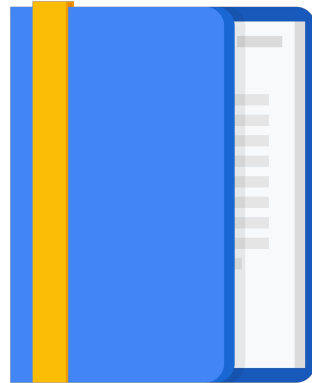
utilizar las alertas de Cloud Monitoring para identificar interrupciones del servicio. Comencemos.

Temario

Administre versiones

Planifique los costos

Supervise los paneles



Comencemos por revisar la administración de versiones.

En una arquitectura de microservicios, es necesario cuidarse de no generar fallas en los clientes cuando se actualizan los servicios

- Incluya la versión en el URI:
 - Si implementa un cambio rotundo, también debe cambiar la versión.
- Es necesario implementar las versiones nuevas sin tiempo de inactividad.
- Es necesario probar las versiones de manera eficaz antes de lanzarlas.

Uno de los beneficios clave de una arquitectura de microservicios es la capacidad de implementarlos de forma independiente. Esto significa que se debe proteger la API del servicio. Se requiere un control de versiones y es importante garantizar la retrocompatibilidad cuando se implementan versiones nuevas. Existen algunas reglas de diseño sencillas que pueden ser útiles, como indicar la versión en el URI y asegurarse de cambiar la versión cuando se aplique un cambio que genere incompatibilidad con versiones anteriores. Implementar nuevas versiones de software siempre conlleva un riesgo. Es importante que nos aseguremos de probarlas de manera eficaz antes de lanzarlas y, cuando estemos listos para implementar una versión nueva, debemos hacerlo sin tiempo de inactividad.

Ahora, veamos algunas estrategias que pueden servir para lograr estos objetivos.

Las actualizaciones progresivas permiten implementar nuevas versiones sin tiempo de inactividad

- Por lo general, hay varias instancias de un servicio detrás de un balanceador de cargas.
 - Las instancias se actualizan una a la vez.
 - Las actualizaciones progresivas son útiles cuando resulta aceptable tener 2 versiones diferentes ejecutándose simultáneamente durante la actualización.
- Las actualizaciones progresivas son una función de los grupos de instancias; simplemente se cambia la plantilla de la instancia.
 - Son la opción predeterminada en Kubernetes; solo se cambia la imagen de Docker.
 - En App Engine esta función está completamente automatizada.

Las actualizaciones progresivas permiten implementar nuevas versiones sin tiempo de inactividad. La configuración típica consiste en tener varias instancias de un servicio tras un balanceador de cargas. De ese modo, una actualización progresiva actualiza una instancia a la vez. Esta estrategia funciona bien si la API no cambia, es retrocompatible, o bien si resulta aceptable tener dos versiones del mismo servicio ejecutándose durante la actualización.

Si usa grupos de instancias, las actualizaciones progresivas son una función integrada. Simplemente se debe definir la estrategia de actualización progresiva cuando la realice.

En Kubernetes, las actualizaciones progresivas están presentes de forma predeterminada. Solo se debe especificar la imagen de Docker de reemplazo. Por último, en App Engine, las actualizaciones progresivas están completamente automatizadas.

Utilice una implementación azul-verde cuando no quiera que se ejecuten al mismo tiempo varias versiones de un servicio

- La implementación azul corresponde a la versión actual.
 - Se crea un entorno totalmente nuevo (el verde).
 - Una vez que se prueba la implementación verde, se migran a ella las solicitudes del cliente.
 - Si se producen errores, se vuelve a la versión anterior.
- En Compute Engine, se puede utilizar DNS para migrar solicitudes de un balanceador de cargas a otro.
 - En Kubernetes, se configura el servicio para enrutar a los nuevos Pods usando etiquetas.
 - Es un cambio sencillo en la configuración.
 - En App Engine, se usa la función División del tráfico.

Utilice una implementación azul-verde cuando no quiera que se ejecuten al mismo tiempo varias versiones de un servicio.

Las implementaciones azul-verde usan dos entornos de implementación completos. La azul ejecuta el software de producción implementado actualmente, mientras que el entorno de la verde está disponible para implementar versiones actualizadas del software.

Cuando se desea probar una nueva versión de software, se la implementa en el entorno verde. Una vez finalizadas las pruebas, la carga de trabajo se pasa del entorno actual (azul) al nuevo (verde). Con esta estrategia, se mitiga el riesgo de una mala implementación, ya que se puede regresar a la implementación anterior si algo sale mal.

En Compute Engine, se puede utilizar DNS para migrar las solicitudes, mientras que en Kubernetes se puede configurar el servicio para enrutar a nuevos Pods usando etiquetas, lo que implica solo un cambio sencillo en la configuración. App Engine permite dividir el tráfico, algo que ya se exploró en el lab anterior de este curso.

Se pueden utilizar lanzamientos Canary antes de una actualización progresiva para reducir el riesgo

- Se sigue ejecutando la versión actual del servicio.
 - Se implementa una instancia de la nueva versión y se le asigna una parte de las solicitudes.
 - Se supervisa para detectar errores.
- En Compute Engine, es posible crear un nuevo grupo de instancias y agregarlo como backend adicional en el balanceador de cargas.
 - En Kubernetes, se crea un nuevo Pod con las mismas etiquetas que los existentes; el servicio enrutará automáticamente hacia él una parte de las solicitudes.
 - En App Engine, se usa la función División del tráfico.

Ahora, puede utilizar lanzamientos Canary antes de una actualización progresiva para reducir el riesgo. Con un lanzamiento Canary, puede realizar una nueva implementación mientras se ejecuta la implementación actual. Después, puede enviar un pequeño porcentaje del tráfico a la nueva implementación y, luego supervisarla.

Cuando se sienta seguro con su nueva implementación, puede enrutar más tráfico a ella, hasta alcanzar el 100%.

En Compute Engine, puede crear un nuevo grupo de instancias y agregarlo al balanceador de cargas como un backend adicional.

En Kubernetes, puede crear un nuevo Pod con las mismas etiquetas que las de los Pods existentes. El servicio desviará automáticamente una parte de las solicitudes al nuevo Pod.

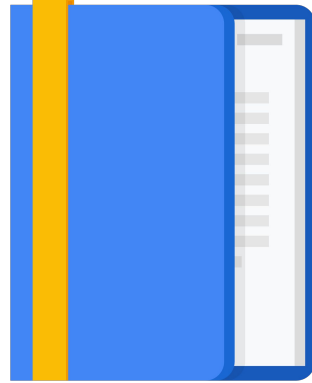
App Engine también permite usar la función de división del tráfico para enviar una parte del tráfico a la nueva versión.

Temario

Administre versiones

Planifique los costos

Supervise los paneles



La planificación de costos es una fase importante del diseño que comienza con la planificación de la capacidad.

La planificación de la capacidad es un ciclo iterativo y continuo

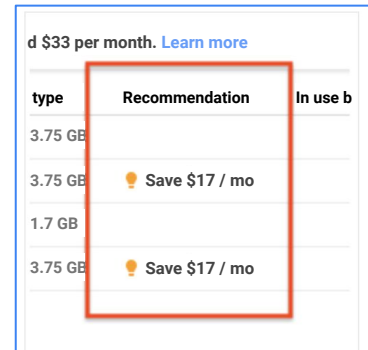


Es recomendable no tratar la planificación de la capacidad como una tarea que se realiza una sola vez, sino como un ciclo iterativo y continuo, como se ilustra en esta diapositiva.

Comience con una previsión en la que se estime la capacidad necesaria. Supervise y revise esta previsión. Luego, asigne en función de los recursos necesarios para alcanzar la capacidad prevista. Esto le permite estimar los costos y equilibrarlos con los riesgos y las recompensas. Cuando se aprueben el diseño y el costo, implemente su diseño y supervíselo para ver qué tan exactos fueron sus previsiones. Estos datos se utilizarán para la siguiente previsión, ya que el proceso se repite.

Optimice el costo de procesamiento

- Comience con VM pequeñas y pruébelas para comprobar si funcionan.
- Considere usar más máquinas pequeñas con el ajuste de escala automático activado.
- Considere los descuentos por compromiso de uso.
- Considere al menos algunas instancias interrumpibles:
 - Se aplica un 80% de descuento.
 - Use la reparación automática para volver a crear VM cuando se interrumpan.
- Las recomendaciones de redimensionamiento de Google Cloud le alertarán cuando las VM no se utilicen mucho.



The screenshot shows a table with VM recommendations. A red box highlights a specific recommendation row. Above the table, there is a text snippet: 'd \$33 per month. [Learn more](#)'. The table has three columns: 'type', 'Recommendation', and 'In use b'. The highlighted row shows a recommendation to change from 3.75 GB to 1.7 GB, saving \$17 per month.

type	Recommendation	In use b
3.75 GB		
3.75 GB	💡 Save \$17 / mo	
1.7 GB		
3.75 GB	💡 Save \$17 / mo	

Un buen punto de partida para quienes buscan optimizar costos es familiarizarse con los precios de las instancias de VM. Por lo general, es bueno comenzar con un par de máquinas pequeñas que puedan escalar horizontalmente mediante el ajuste de escala automático a medida que aumente la demanda.

Para optimizar el costo de sus máquinas virtuales, considere utilizar descuentos por compromiso de uso, ya que estos pueden ser significativos. Además, si sus cargas de trabajo admiten instancias interrumpibles, puede ahorrar hasta un 80% y utilizar la reparación automática para recuperar las instancias que se interrumpan.

Compute Engine también proporciona recomendaciones de redimensionamiento para sus instancias de VM, como se muestra a la derecha. Esta es una función muy útil con la que puede seleccionar el tamaño de VM correcto para sus cargas de trabajo y optimizar costos.

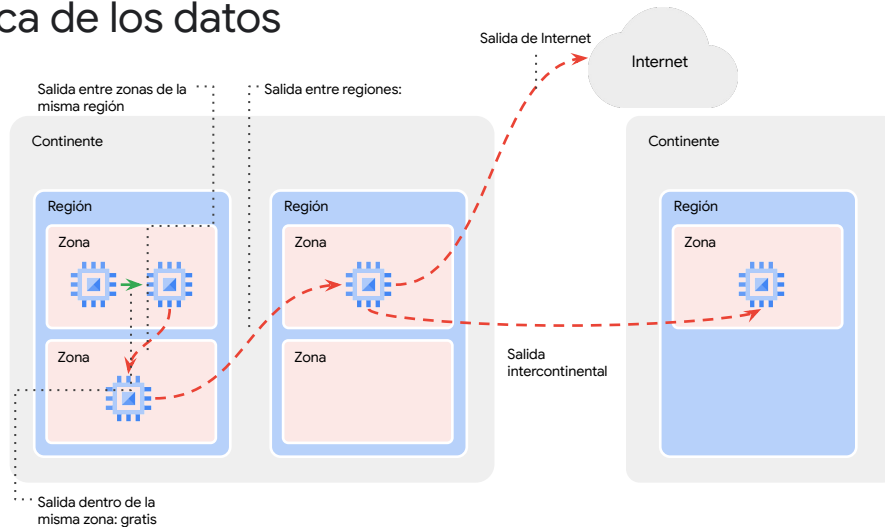
Optimice el costo del disco

- No sobreasigne el espacio en el disco.
- Determine cuáles son las características de rendimiento que requieren sus aplicaciones:
 - Patrón de E/S: operaciones de lectura y escritura pequeñas o grandes
 - Configure sus instancias para optimizar el rendimiento del almacenamiento.
- Según los requisitos de E/S, considere usar discos estándar en lugar de SSD.

Capacidad mensual	PD estándar	PD SSD
10 GB	\$0.40	\$1.70
1 TB	\$40	\$170
16 TB	\$655.36	\$5,570.56

Un error común es sobreasignar espacio en el disco. Esto no es rentable. Sin embargo, hay otros factores, aparte del tamaño del disco, que deben tenerse en cuenta. Es importante determinar el rendimiento de sus aplicaciones: los patrones de E/S o si tienen operaciones de escritura y lectura grandes o pequeñas, o bien usan principalmente datos de solo lectura. Con este tipo de información, podrá seleccionar el tipo de disco correcto. Como se muestra en la tabla, los discos persistentes SSD son mucho más caros que los discos persistentes estándar. Comprender sus patrones de E/S le permite ahorrar sumas significativas.

Para optimizar los costos de red, mantenga las máquinas cerca de los datos

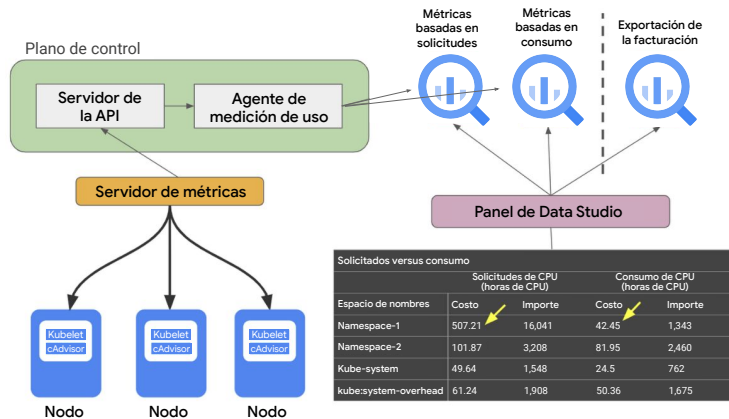


Para optimizar los costos de red, la práctica recomendada es mantener las máquinas lo más cerca posible de los datos a los que deben acceder. En este gráfico, se muestran los distintos tipos de salida: dentro de la misma zona, entre zonas de la misma región, intercontinental y de Internet. Es importante tener en cuenta los costos de salida. No son todos iguales. La salida dentro de la misma zona es gratis. La salida hacia otro servicio de Google Cloud dentro de la misma región con una dirección IP interna o externa es gratuita, excepto para algunos servicios como Memorystore o Redis. La salida entre zonas de la misma región se cobra, así como la salida de Internet.

Una forma de optimizar sus costos de red es mantener las máquinas cerca de los datos.

Medir el uso de GKE puede prevenir el aprovisionamiento excesivo de los clústeres de Kubernetes

Se comparan los recursos solicitados con los consumidos.



Otra forma de optimizar el costo es aprovechar la medición de uso de GKE, con la que se puede prevenir el aprovisionamiento excesivo en los clústeres de Kubernetes.

Con la medición de uso de GKE, un agente recopila las métricas de consumo y las solicitudes de recursos mediante un sondeo de objetos PodMetrics del servidor de métricas. Los registros de solicitudes y consumo de recursos se exportan a dos tablas diferentes en un conjunto de datos de BigQuery que especifica el usuario. Cuando se comparan los recursos solicitados y los consumidos, resulta fácil detectar y corregir posibles desperdicios.

En este gráfico, se muestra una configuración típica en la que se usa BigQuery para las métricas basadas en solicitudes que recopila el agente de medición de uso, que se analizan en Data Studio junto con los datos obtenidos de la exportación de facturación.

Compare los costos de las distintas alternativas de almacenamiento antes de decidirse por una o la otra

Elija un servicio de almacenamiento que satisfaga sus requisitos de capacidad por un costo razonable:

- Almacenar 1 GB en Firestore es gratis.
- Almacenar 1 GB en Cloud Bigtable puede costar alrededor de \$1,400 por mes.

Anteriormente en el curso, hablamos sobre los distintos servicios de almacenamiento. Es importante comparar los costos y las características de las diferentes opciones.

Por ejemplo, en el momento en que estoy haciendo esta grabación, Firestore proporciona 1 GB de almacenamiento gratis con el nivel de acceso gratuito. Si almacenáramos esa misma cantidad de datos en Cloud Bigtable, podríamos terminar pagando más de \$1,000 por mes, ya que se necesitarían al menos 3 nodos de Bigtable.

En otras palabras, el servicio de almacenamiento y bases de datos elegido puede influir de manera significativa en su factura.

Compare los costos de las distintas alternativas de almacenamiento antes de decidirse por una o la otra

Elija un servicio de almacenamiento que satisfaga sus requisitos de capacidad por un costo razonable:

- Almacenar 1 GB en Firestore es gratis.
- Almacenar 1 GB en Cloud Bigtable puede costar alrededor de \$1,400 por mes.

Anteriormente en el curso, hablamos sobre los distintos servicios de almacenamiento.

Es importante comparar los costos y las características de las diferentes opciones.

Compare los costos de las distintas alternativas de almacenamiento antes de decidirse por una o la otra

Elija un servicio de almacenamiento que satisfaga sus requisitos de capacidad por un costo razonable:

- Almacenar 1 GB en Firestore es gratis.
- Almacenar 1 GB en Cloud Bigtable puede costar alrededor de \$1,400 por mes.

Por ejemplo, en el momento en que estoy haciendo esta grabación, Firestore proporciona 1 GB de almacenamiento gratis con el nivel de acceso gratuito.

Compare los costos de las distintas alternativas de almacenamiento antes de decidirse por una o la otra

Elija un servicio de almacenamiento que satisfaga sus requisitos de capacidad por un costo razonable:

- Almacenar 1 GB en Firestore es gratis.
- Almacenar 1 GB en Cloud Bigtable puede costar alrededor de \$1,400 por mes.

Si almacenáramos esa misma cantidad de datos en Cloud Bigtable, podríamos terminar pagando más de \$1,000 por mes, ya que se necesitarían al menos 3 nodos de Bigtable.

Compare los costos de las distintas alternativas de almacenamiento antes de decidirse por una o la otra

Elija un servicio de almacenamiento que satisfaga sus requisitos de capacidad por un costo razonable:

- Almacenar 1 GB en Firestore es gratis.
- Almacenar 1 GB en Cloud Bigtable puede costar alrededor de \$1,400 por mes.

En otras palabras, el servicio de almacenamiento y bases de datos elegido puede influir de manera significativa en su factura.

Considere servicios alternativos para ahorrar costos en lugar de asignar más recursos

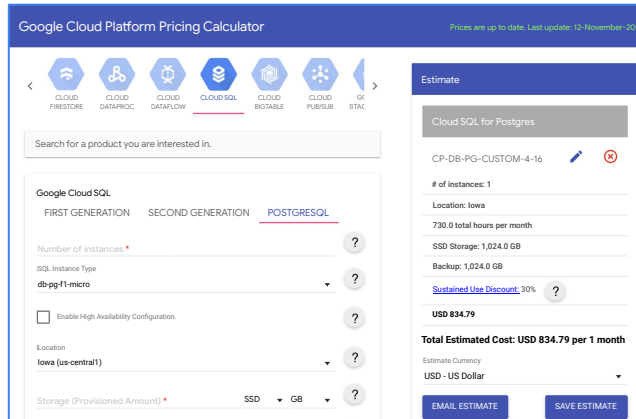
- CDN
- Almacenar en caché
- Mensajes
- Agregar a una cola
- Etcétera

El diseño de la arquitectura también puede ayudar a optimizar los costos.

Por ejemplo, si usa Cloud CDN para contenido estático o Memorystore como una caché, puede ahorrar en lugar de asignar más recursos. Asimismo, en vez de utilizar un almacén de datos entre dos aplicaciones, considere enviar mensajes o agregar a una cola con Pub/Sub para separar los servicios de comunicación y reducir las necesidades de almacenamiento.

Utilice la calculadora de precios de Google Cloud para estimar los costos

- Base las estimaciones de sus costos en las previsiones y la planificación de la capacidad.
- Compare los costos de distintos servicios de procesamiento y almacenamiento.

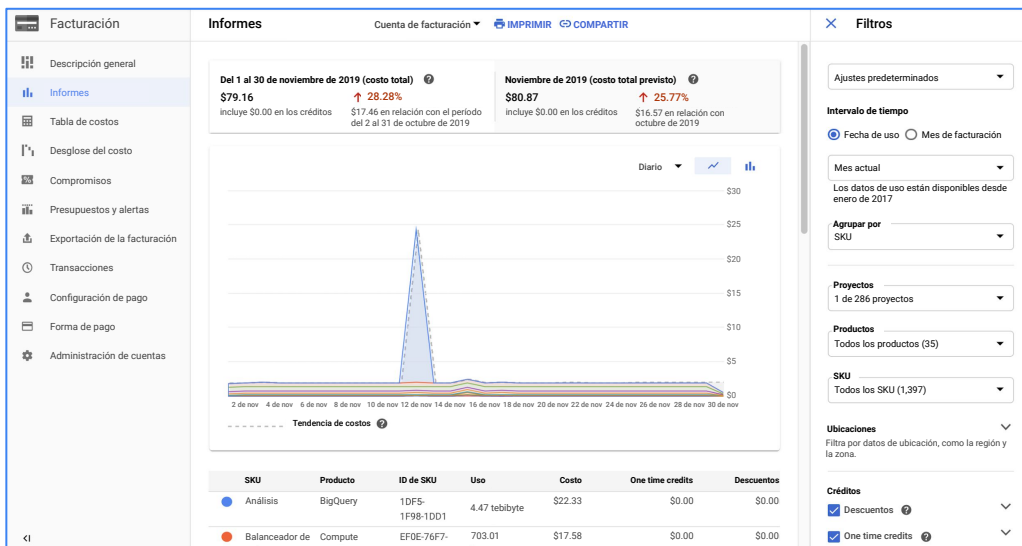


The screenshot displays the Google Cloud Platform Pricing Calculator interface. The top navigation bar includes icons for various services: Cloud Firestore, Cloud DataProc, Cloud Dataflow, Cloud SQL (selected), Cloud Bigtable, Cloud Pub/Sub, and GKE. The main content area is titled 'Google Cloud SQL' and features tabs for 'FIRST GENERATION', 'SECOND GENERATION', and 'POSTGRES SQL' (which is active). Below these tabs, there are input fields for 'Number of Instances' (set to 1), 'SQL Instance Type' (set to db-pg11-micro), 'Location' (set to Iowa (us-central1)), and 'Storage (Provisioned Amount)' (set to 1,024.0 GB). A 'Sustained Use Discount' of 30% is indicated. The right sidebar, titled 'Estimate', shows the configuration details: 'CP-DB-PG-CUSTOM-4-16', '1 instance', 'Iowa', '730.0 total hours per month', '1,024.0 GB SSD Storage', and '1,024.0 GB Backup'. The 'Total Estimated Cost' is displayed as 'USD 834.79 per 1 month'. At the bottom of the sidebar, there are buttons for 'EMAIL ESTIMATE' and 'SAVE ESTIMATE'.

<https://cloud.google.com/products/calculator>

La calculadora de precios debe ser su recurso principal para estimar los costos. Estos deben basarse en sus previsiones y la planificación de la capacidad. Esta herramienta es excelente para comparar precios de distintos servicios de procesamiento y almacenamiento, y lo utilizará en la actividad de diseño que realizaremos más adelante.

Los informes de facturación proporcionan desgloses detallados del costo



Para supervisar los costos de su servicio existente, aproveche la página Informes de Facturación de Cloud como se muestra aquí. En este informe, se muestran los cambios de los costos en comparación con el mes anterior, y puede usar los filtros para buscar proyectos, regiones y productos específicos, como se observa a la derecha.

Este informe también contiene las recomendaciones de redimensionamiento para sus instancias de Compute Engine.

Para realizar un análisis avanzado de los costos, exporte los datos de facturación a BigQuery

Facturación

Exportación de la facturación Cuenta de facturación de proyectos de ROIGCP Instructor

BIGQUERY EXPORT EXPORTAR ARCHIVOS

BigQuery Export envía tus datos de facturación a un conjunto de datos de BigQuery. [Más información sobre BigQuery](#)

Detalle del costo diario

✓ **Habilitado**

Cada día, se actualizará el conjunto de datos de BigQuery seleccionado con el detalle diario de tu costo por SKU.

Nombre del proyecto roigcp-operations

Nombre del conjunto de datos [ROIGCPBilling](#)

[EDITAR CONFIGURACIÓN](#) [INHABILITAR BIGQUERY EXPORT](#)

Para los análisis de costos avanzados, recomiendo exportar los datos de facturación a BigQuery, como se muestra en esta captura de pantalla. De este modo, podrá analizar los datos de facturación para identificar grandes gastos y optimizar sus inversiones en Google Cloud.

Por ejemplo, supongamos que etiqueta las instancias de VM que se encuentran en diferentes regiones. Estas instancias podrían estar enviando la mayor parte del tráfico a otro continente, lo cual generaría costos más altos. En ese caso, considere reubicar algunas de esas instancias o usar un servicio de almacenamiento en caché como Cloud CDN para almacenar en caché el contenido más cercano a sus usuarios, lo que disminuye la inversión en herramientas de redes.

Visualice la inversión con Google Data Studio



Panel de facturación



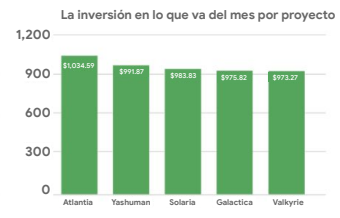
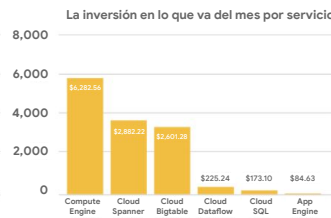
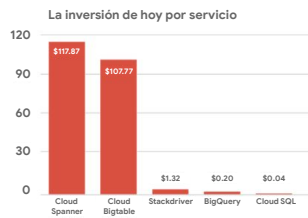
Vista diaria



Vista mensual



General



Incluso puede visualizar la inversión en el tiempo con Google Data Studio, que convierte los datos en informes y paneles informativos totalmente personalizables y fáciles de leer y compartir.

El servicio de datos se muestra en una vista diaria y en una mensual con resúmenes que también se pueden desglosar para obtener estadísticas más significativas.

Configure presupuestos y alertas para que su equipo sepa cuánto invierte

Nombre*	Porcentaje del presupuesto	Importe	Cuándo se activa ?
<input type="text" value="Nombre de presupuesto"/>	<input type="text" value="50"/> %	<input type="text" value="\$ 250"/>	<input type="text" value="Real"/>
Proyectos <input type="text" value="Todos los proyectos (2)"/>	<input type="text" value="90"/> %	<input type="text" value="\$ 450"/>	<input type="text" value="Real"/>
	<input type="text" value="100"/> %	<input type="text" value="\$ 500"/>	<input type="text" value="Real"/>

2 Importe

Tipo de presupuesto

[Importe especificado](#)

Un importe fijo con el que se comparará tu inversión.

Importe objetivo

Inversión del último mes

Presupuestos programáticos: Cloud Pub/Sub → Cloud Functions

Puede establecer un presupuesto para facilitar la planificación de proyectos y el control de costos. Cuando se configura un presupuesto, es posible hacer un seguimiento de la inversión con respecto a ese importe. En esta captura de pantalla, se muestra la interfaz para crear presupuestos:

1. Establezca un nombre de presupuesto y especifique a qué proyecto se aplica.
2. Establezca el presupuesto en un importe específico o haga que coincida con la inversión del último mes.
3. Configure las alertas de presupuesto, que envían correos electrónicos a los administradores de facturación una vez que la inversión supera un porcentaje del presupuesto o un importe específico.

En este caso, el correo electrónico se enviaría cuando la inversión alcance el 50%, el 90% y el 100% del importe del presupuesto. Incluso puede elegir que se envíe una alerta cuando se prevea que la inversión superará el porcentaje del importe del presupuesto al final del período presupuestado.

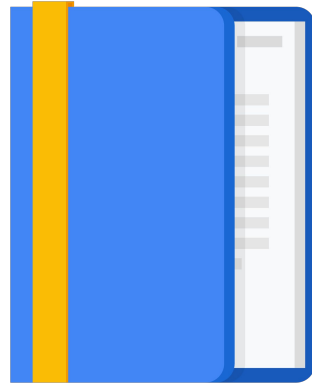
Además de recibir un correo electrónico, puede usar las notificaciones de Pub/Sub para recibir de manera programática actualizaciones de la inversión en relación con este presupuesto. Incluso puede crear una Cloud Function que escuche el tema de Pub/Sub para automatizar la administración del costo.

Temario

Administre versiones

Planifique los costos

Supervise los paneles



Ahora, veamos cómo supervisar y visualizar información con paneles.

Google Cloud unifica las herramientas que necesita para supervisar los SLO y ANS de su servicio

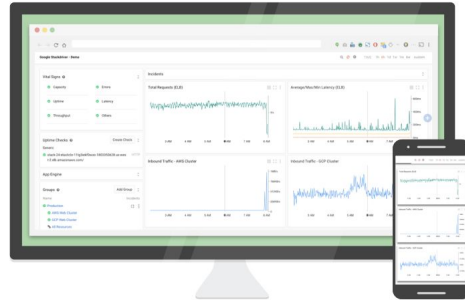


Google Cloud unifica las herramientas que necesita para supervisar los SLO y ANS de su servicio en tiempo real.

Entre ellas, se incluyen Monitoring, Logging, Trace, Debugger, Error Reporting y Profiler. Todas estas herramientas permiten obtener las estadísticas necesarias para lograr sus SLO y determinar el motivo principal en esos casos poco comunes en los que no los alcanza.

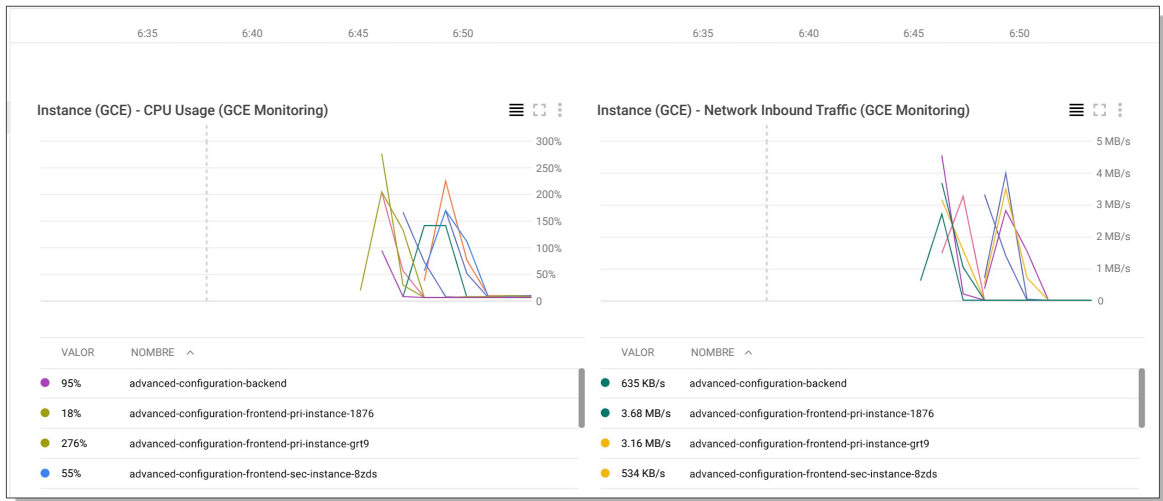
Los paneles de Monitoring supervisan sus servicios

- Supervise lo que paga por:
 - Uso de CPU
 - Capacidad de almacenamiento
 - Lecturas y escrituras
 - Salida de red
 - Etcétera
- Supervise sus SLI para determinar si está logrando sus SLO.



Los paneles son una forma de ver y analizar los datos de las métricas que son importantes para usted. Esto incluye sus SLI para asegurarse de que está cumpliendo con los ANS. En la página Monitoring de Cloud Console, se proporcionan automáticamente paneles predefinidos sobre los recursos y servicios que utiliza. Es importante que supervise lo que paga para determinar tendencias, cuellos de botella y ahorros potenciales.

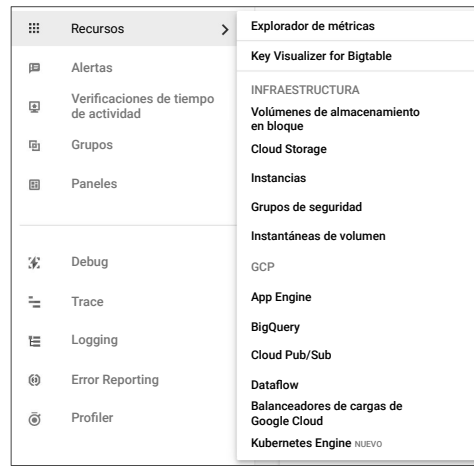
Gráficos de ejemplo en un panel de Monitoring



Este es un ejemplo de algunos gráficos en un panel de Monitoring. En la parte izquierda, puede ver el uso de CPU de varias instancias de Compute Engine y, en la parte derecha, se muestra el tráfico de entrada de cada una.

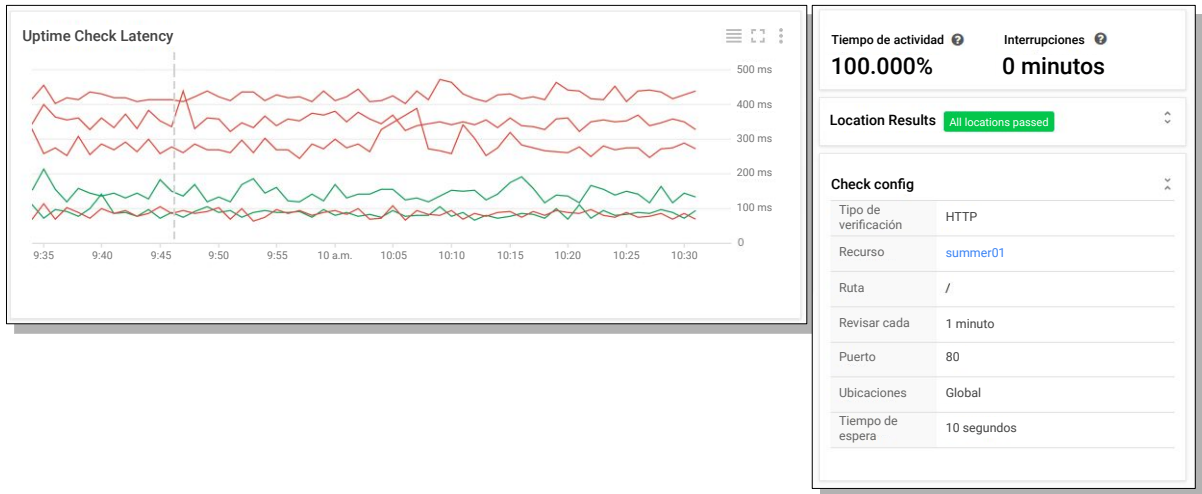
Este tipo de gráfico proporciona información valiosa sobre los patrones de uso.

Para comenzar, Cloud Monitoring crea paneles predeterminados sobre los recursos de sus proyectos



Para comenzar, Cloud Monitoring crea paneles predeterminados sobre los recursos de sus proyectos, como se muestra en esta captura de pantalla. También puede crear paneles personalizados, que podrá explorar en el próximo lab.

Cree verificaciones de tiempo de actividad para supervisar la disponibilidad y la latencia



Supervisar la latencia es una buena idea, porque con ello se pueden detectar los problemas rápidamente antes de que ocurran. Como se muestra en esta diapositiva, puede crear fácilmente verificaciones de tiempo de actividad para supervisar la disponibilidad y la latencia de sus servicios. Hasta ahora, hay un 100% de tiempo de actividad sin interrupciones.

La latencia es uno de los cuatro indicadores clave que se mencionan en el libro de ingeniería de confiabilidad de sitios, o SRE, de Google. La SRE es una disciplina que aplica diferentes aspectos de la ingeniería de software a operaciones cuyo objetivo es crear sistemas de software ultraescalables y altamente confiables. Esta disciplina le ha permitido a Google crear, implementar, supervisar y mantener algunos de los sistemas de software más grandes del mundo.

Agregué el vínculo al libro de SRE a las diapositivas de este módulo [<https://landing.google.com/sre/books/>].

Cree alertas cuando su servicio no cumpla los SLO

Create new alerting policy

1 Condiciones

Basic Conditions

HTTP check on instance summer01

Violates when: Uptime Check Health on Instance (GCE) summer01 fails

[Editar](#) [Borrar](#)

+

 Agregar otra condición

2 Notifications (optional)

When alerting policy violations occur, you will be notified via these channels. [Más información](#)

Correo electrónico

demo@example.com

×

+

 Add Another Notification

3 Documentation (optional)

When email notifications are sent, they'll include any text entered here. This can convey useful information about the problem and ways to approach fixing it.

Editar

 Vista previa [Markdown Formatting Help](#)

Falló la verificación de estado del servidor principal
+ El servidor con el nombre summer01 falló una verificación de tiempo de actividad de Stackdriver
+ La dirección IP del servidor es la siguiente: 104.197.58.79

4 Name this policy

A policy's name is used in identifying which policies were triggered, as well as managing configurations of different policies.

Política de verificación de tiempo de actividad

Guardar política

 Cancelar

Su SLO será más estricto que su ANS, así que es importante recibir alertas cuando no se cumple un SLO, ya que eso es una advertencia temprana de que el ANS está en peligro.

Este es un ejemplo de cómo luce el proceso para crear una política de alertas. A la izquierda, puede ver una condición de verificación de HTTP en la instancia summer01. Esta condición enviará un correo electrónico que se personaliza con el contenido de la sección de documentación a la derecha.

Actividad 13: Estime y planifique los costos

Consulte el Cuaderno de ejercicios de Design and Process

- Utilice la calculadora de precios para crear una estimación inicial a fin de implementar la aplicación del caso de éxito.



En esta actividad de diseño, use la calculadora de precios de Google Cloud para crear una estimación inicial a fin de implementar la aplicación del caso de éxito.

Actividad 13: Estime y planifique los costos

Consulte el Cuaderno de ejercicios de Design and Process

- Utilice la calculadora de precios para crear una estimación inicial a fin de implementar la aplicación del caso de éxito.



En esta actividad de diseño,

Actividad 13: Estime y planifique los costos

Consulte el Cuaderno de ejercicios de Design and Process

- Utilice la calculadora de precios para crear una estimación inicial a fin de implementar la aplicación del caso de éxito.



utilice la calculadora de precios de Google Cloud para crear una estimación inicial a fin de implementar la aplicación del caso de éxito.

Estimate

Cloud SQL for Postgres

CP-DB-PG-CUSTOM-4-16

of instances: 1

Location: Iowa

730.0 total hours per month

SSD Storage: 500.0 GB

Backup: 0.0 GB

[Sustained Use Discount](#): 30%

USD 574.71

Total Estimated Cost: USD 574.71 per 1 month

Estimate Currency

USD - US Dollar

EMAIL ESTIMATE

SAVE ESTIMATE

La calculadora de precios le proporciona un formulario para cada servicio, y puede completar uno de ellos a fin de estimar el costo de uso de un servicio determinado. Por ejemplo, en esta captura de pantalla calculé el costo de una instancia de SQL personalizada con 4 núcleos, 16 GB de RAM y 500 GB de almacenamiento SSD. Esto podría representar la base de datos de pedidos de mi aplicación de viajes en línea.

Algunas de estas estimaciones no son fáciles de generar, ya que posiblemente no sepamos cuántos datos se necesitan en los servicios de almacenamiento y bases de datos ni cuánta capacidad de procesamiento requieren las plataformas de implementación. Sin embargo, puede ser más difícil estimar elementos como la salida de red o la cantidad de operaciones de lectura y escritura. Comience con una estimación aproximada y defínala mejor a medida que sus planes de capacidad mejoren.

Consulte la actividad 13 en su cuaderno de ejercicios a fin de ver estimaciones de costo similares para el caso de éxito.

Revisión de la actividad 13: Estime y planifique los costos

- Utilice la calculadora de precios para crear una estimación inicial a fin de implementar la aplicación del caso de éxito.



En esta actividad, se le solicitó usar la calculadora de precios de Google Cloud para estimar el costo de la aplicación del caso de éxito.

Nombre del servicio	Recurso de Google Cloud	Costo mensual
Pedidos	Cloud SQL	\$1,264.44
Inventario	Firestore	\$215.41
Inventario	Cloud Storage	\$1,801.00
Analytics	BigQuery	\$214.72

Esta es una estimación aproximada para las aplicaciones de base de datos de TurisClic, nuestro portal de viajes en línea.

Ajusté la base de datos de pedidos a fin de incluir una réplica de conmutación por error a fin de lograr una alta disponibilidad y elaboré algunas estimaciones generales para mis otros servicios. El servicio de inventario utiliza Cloud Storage para almacenar datos JSON guardados en archivos de texto. Dado que este es el servicio más costoso, debería reconsiderar la clase de almacenamiento o configurar la administración del ciclo de vida de los objetos.

Este es solo un ejemplo. Los costos dependerán de su caso de éxito.

Lab

Supervise aplicaciones en Google Cloud



Objetivos

- Examinar los registros de Cloud
- Ver la información de Profiler
- Explorar Cloud Trace
- Supervisar los recursos mediante paneles
- Crear alertas y verificaciones de tiempo de actividad

Comenzamos este curso definiendo los SLO y SLI de sus servicios. Esto ayuda a crear un diseño y una arquitectura detallados, y permite que los desarrolladores sepan cuándo terminaron de implementar un servicio.

Sin embargo, los SLI y SLO no son muy útiles si no supervisa sus aplicaciones para comprobar que se los alcance. Es por ello que debemos utilizar herramientas de supervisión. En este lab, aprenderá a usar algunas de ellas.

Específicamente, examinará registros, visualizará información de Profiler, supervisará sus recursos y explorará seguimientos con paneles, y creará alertas y verificaciones de tiempo de actividad.

Revisión del lab

Supervise aplicaciones en Google Cloud

En este lab, vio cómo supervisar sus aplicaciones con herramientas integradas de Google Cloud. Primero, implementó una aplicación en App Engine y examinó registros de Cloud. Después, visualizó información de Profiler y exploró Cloud Trace. Finalmente, supervisó su aplicación con paneles y creó alertas y verificaciones de tiempo de actividad.

Puede continuar con un recorrido por el lab, pero recuerde que la interfaz de usuario de Google Cloud puede cambiar, por lo que su entorno podría ser un poco diferente.

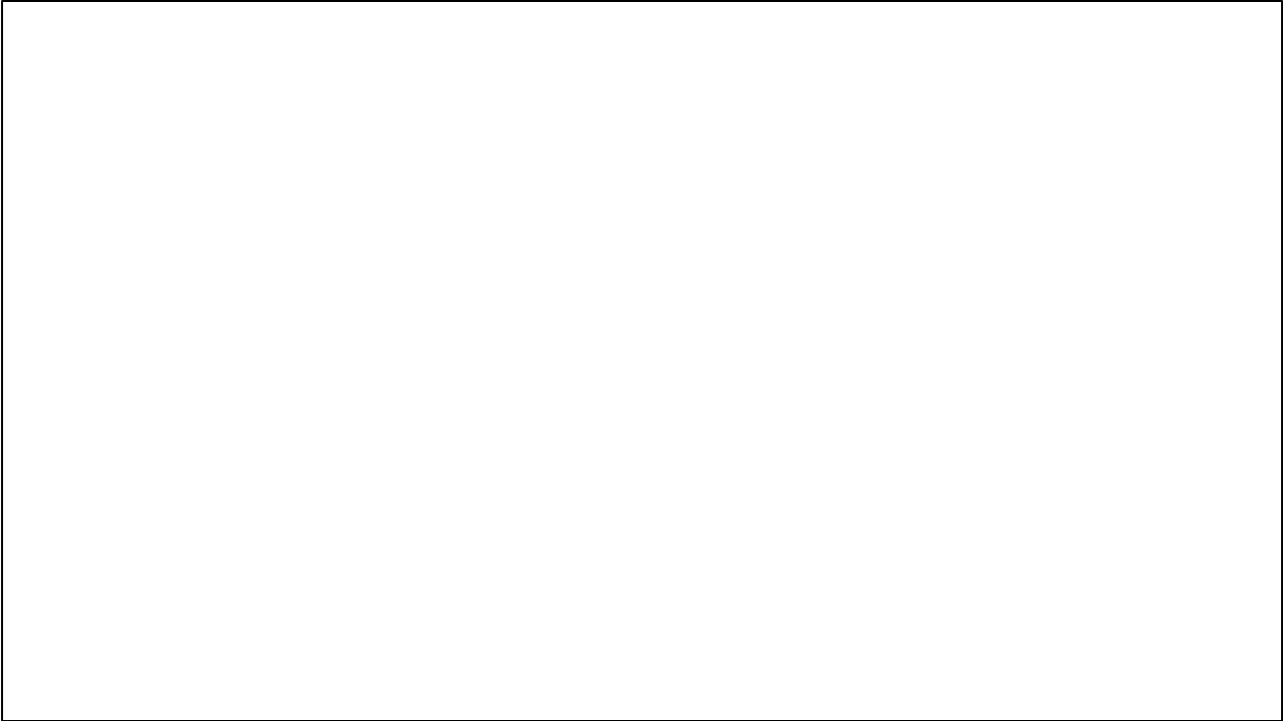
Repaso

Mantenimiento y supervisión

En este módulo, aprendió a administrar versiones nuevas de sus microservicios usando actualizaciones progresivas, implementaciones de versiones canary y, además, implementaciones azul-verde. Cuando se trabaja con microservicios, es importante que las versiones nuevas se implementen sin tiempo de inactividad y que no ocasionen errores en los clientes que usan los servicios.

También aprendió sobre la planificación de costos y la optimización, y estimó el costo de ejecutar la aplicación del caso de éxito.

Al final del módulo, aprendió a aprovechar las herramientas de supervisión proporcionadas por Google Cloud. Estas pueden ser muy valiosas para administrar sus servicios y supervisar sus SLI y SLO.



[P] Gracias por tomar el curso “Reliable Cloud Infrastructure: Design and Process”. Esperamos que ahora comprenda mejor cómo diseñar aplicaciones y servicios que aprovechen óptimamente los servicios de la plataforma que ofrece Google Cloud.

[S] Además, esperamos que las actividades de diseño y los labs lo ayuden a sentirse más cómodo con el diseño y el proceso en Google Cloud.

[P] Ahora es su turno. Aplique lo que aprendió para diseñar sus propias aplicaciones, implementaciones y supervisión.

[S] Hasta la próxima.

Cuestionario

Suponga que realizó una corrección menor en uno de sus servicios y quiere implementar la versión nueva sin tiempo de inactividad. ¿Cuál operación elegiría?

- A. Actualización progresiva
- B. Implementación azul-verde
- C. Implementación de versiones canary
- D. Prueba A/B

Suponga que realizó una corrección menor en uno de sus servicios y quiere implementar la versión nueva sin tiempo de inactividad. ¿Cuál operación elegiría?

- A. Actualización progresiva
- B. Implementación azul-verde
- C. Implementación de versiones canary
- D. Prueba A/B

Cuestionario

Suponga que realizó una corrección menor en uno de sus servicios y quiere implementar la versión nueva sin tiempo de inactividad. ¿Cuál operación elegiría?

A. Actualización progresiva

B. Implementación azul-verde

C. Implementación de versiones canary

D. Prueba A/B

- A. Esta respuesta es correcta. Una actualización progresiva actualiza las instancias de forma incremental hasta que se actualizan todas. Esta estrategia se puede controlar, por ejemplo, mediante un grupo de instancias administrado con Compute Engine o GKE.
- B. Esta opción no es correcta. Para una implementación azul-verde, se requieren dos implementaciones completas. Según la pregunta, se necesita actualizar un servicio específico, no toda la aplicación.
- C. Esta opción no es correcta. Esta es una posible solución para un servicio que busque eliminar o reducir los riesgos mediante la aplicación de la actualización a un subconjunto pequeño de usuarios, de modo que estos prueben una función nueva. En este caso, estamos hablando de una corrección menor, por lo que sería mejor hacer una actualización progresiva.
- D. Esta opción no es correcta. Las pruebas A/B se usan generalmente para comparar los patrones de comportamiento/utilización de dos versiones de una aplicación o página web.

La respuesta A es correcta. Una actualización progresiva actualiza las instancias de forma incremental hasta que se actualizan todas. Esta estrategia se puede controlar, por ejemplo, mediante un grupo de instancias administrado con Compute Engine o GKE.

La respuesta B no es correcta. Para una implementación azul-verde, se requieren dos implementaciones completas. Según la pregunta, se necesita actualizar un servicio específico, no toda la aplicación.

La respuesta C no es correcta. Esta es una posible solución para un servicio que busque eliminar o reducir los riesgos mediante la aplicación de la actualización a un subconjunto pequeño de usuarios, de modo que estos prueben una función nueva. En este caso, estamos hablando de una corrección menor, por lo que sería mejor hacer una actualización progresiva.

La respuesta D no es correcta. Las pruebas A/B se usan generalmente para comparar los patrones de comportamiento/utilización de dos versiones de una aplicación o página web.

Cuestionario

Suponga que realizó una actualización menor de un servicio. Para probarla en producción, quiere enviar una pequeña parte de las solicitudes a la versión nueva. ¿Cuál operación elegiría?

- A. Actualización progresiva
- B. Implementación azul-verde
- C. Implementación de versiones canary
- D. Prueba A/B

Suponga que realizó una actualización menor de un servicio. Para probarla en producción, quiere enviar una pequeña parte de las solicitudes a la versión nueva. ¿Cuál operación elegiría?

- A. Actualización progresiva
- B. Implementación azul-verde
- C. Implementación de versiones canary
- D. Prueba A/B

Cuestionario

Suponga que realizó una actualización menor de un servicio. Para probarla en producción, quiere enviar una pequeña parte de las solicitudes a la versión nueva. ¿Cuál operación elegiría?

- A. Actualización progresiva
- B. Implementación azul-verde
- C. Implementación de versiones canary
- D. Prueba A/B

- A. Esta opción no es correcta. Una actualización progresiva actualiza las instancias de forma incremental hasta que se actualizan todas. Esto no es lo que se pide en la pregunta (enviar una pequeña parte de las solicitudes).
- B. Esta opción no es correcta. Para una implementación azul-verde, se requieren dos implementaciones completas. Según la pregunta, se necesita enviar una pequeña parte de las solicitudes a un servicio actualizado.
- C. Esta respuesta es correcta. Las implementaciones de versiones canary tienen el propósito de eliminar o reducir los riesgos mediante la aplicación de la actualización a un pequeño subconjunto de usuarios, de modo que estos prueben una función nueva. Es la opción que mejor se ajusta a lo que se pide en la pregunta.
- D. Esta opción no es correcta. Las pruebas A/B se usan generalmente para comparar los patrones de comportamiento/utilización de dos versiones de una aplicación o página web.

La respuesta C es correcta. Las implementaciones de versiones canary tienen el propósito de eliminar o reducir los riesgos mediante la aplicación de la actualización a un pequeño subconjunto de usuarios, de modo que estos prueben una función nueva. Es la opción que mejor se ajusta a lo que se pide en la pregunta.

La respuesta A no es correcta. Una actualización progresiva actualiza las instancias de forma incremental hasta que se actualizan todas. Esto no es lo que se pide en la pregunta (enviar una pequeña parte de las solicitudes).

La respuesta B no es correcta. Para una implementación azul-verde, se requieren dos implementaciones completas. Según la pregunta, se necesita enviar una pequeña parte de las solicitudes a un servicio actualizado.

La respuesta D no es correcta. Las pruebas A/B se usan generalmente para comparar los patrones de comportamiento/utilización de dos versiones de una aplicación o página web.

Cuestionario

Suponga que va a implementar entornos de pruebas mediante VM de Compute Engine. Si bien puede aceptar un poco de tiempo de inactividad, es muy importante que las implementaciones sean tan económicas como sea posible. ¿Cuál de las siguientes opciones puede ayudarlo a ahorrar la mayor cantidad de dinero?

- A. Descuento por compromiso de uso
- B. Descuento por uso continuo
- C. Nodos de usuario único
- D. Máquinas interrumpibles

Suponga que va a implementar entornos de pruebas mediante VM de Compute Engine. Si bien puede aceptar un poco de tiempo de inactividad, es muy importante que las implementaciones sean tan económicas como sea posible. ¿Cuál de las siguientes opciones puede ayudarlo a ahorrar la mayor cantidad de dinero?

- A. Descuento por compromiso de uso
- B. Descuento por uso continuo
- C. Nodos de usuario único
- D. Máquinas interrumpibles

Cuestionario

Suponga que va a implementar entornos de pruebas mediante VM de Compute Engine. Si bien puede aceptar un poco de tiempo de inactividad, es muy importante que las implementaciones sean tan económicas como sea posible. ¿Cuál de las siguientes opciones puede ayudarlo a ahorrar la mayor cantidad de dinero?

- A. Descuento por compromiso de uso
- B. Descuento por uso continuo
- C. Nodos de usuario único
- D. Máquinas interrumpibles

- A. Esta opción no es correcta. El compromiso de uso requiere un plazo de 1 o 3 años. Dado que usted va a implementar entornos de pruebas y puede tolerar un poco de tiempo de inactividad, esta no es la solución más rentable.
- B. Esta opción no es correcta. Los descuentos por uso continuo se aplican automáticamente cuando se pasa determinado umbral. Dado que usted va a implementar entornos de pruebas, es poco probable que se alcancen los umbrales mensuales requeridos para obtener descuentos, por lo que esta no es la solución más rentable.
- C. Esta opción no es correcta. Los nodos de usuario único no son la solución más rentable para este caso de uso.
- D. Esta respuesta es correcta. Dados los requisitos que se mencionan en la pregunta (lograr el costo más bajo, incluso si hay un poco de tiempo de inactividad), las máquinas interrumpibles son la solución más rentable, ya que tienen un costo hasta un 80% menor que las máquinas no interrumpibles.

La respuesta D es correcta. Dados los requisitos que se mencionan en la pregunta (lograr el costo más bajo, incluso si hay un poco de tiempo de inactividad), las máquinas interrumpibles son la solución más rentable, ya que tienen un costo hasta un 80% menor que las máquinas no interrumpibles.

La respuesta A no es correcta. El compromiso de uso requiere un plazo de 1 o 3 años. Dado que usted va a implementar entornos de pruebas y puede tolerar un poco de tiempo de inactividad, esta no es la solución más rentable.

La respuesta B no es correcta. Los descuentos por uso continuo se aplican automáticamente cuando se pasa determinado umbral. Dado que usted va a implementar entornos de pruebas, es poco probable que se alcancen los umbrales mensuales requeridos para obtener descuentos, por lo que esta no es la solución más rentable.

La respuesta C no es correcta. Los nodos de usuario único no son la solución más rentable para este caso de uso.

Cuestionario

Suponga que tiene un servicio con un SLO de disponibilidad del 99%.
¿Qué herramienta puede utilizar para supervisar si está cumpliendo con ese porcentaje?

- A. Verificación de estado
- B. Verificación del tiempo de actividad
- C. Sondeo de funcionamiento
- D. Sondeo de preparación

Suponga que tiene un servicio con un SLO de disponibilidad del 99%. ¿Qué herramienta puede utilizar para supervisar si está cumpliendo con ese porcentaje?

- A. Verificación de estado
- B. Verificación del tiempo de actividad
- C. Sondeo de funcionamiento
- D. Sondeo de preparación

Cuestionario

Suponga que tiene un servicio con un SLO de disponibilidad del 99%.
¿Qué herramienta puede utilizar para supervisar si está cumpliendo con ese porcentaje?

- A. Verificación de estado
- B. Verificación del tiempo de actividad**
- C. Sondeo de funcionamiento
- D. Sondeo de preparación

B. Esta respuesta es correcta. La disponibilidad es el porcentaje de tiempo que un sistema está en ejecución y puede procesar solicitudes. Cuando se supervisa esa métrica, se habilita la derivación de la métrica de SLO.

La verificación de estado no es la respuesta correcta, ya que determina si la aplicación está disponible en un momento determinado, pero no cuánto tiempo lleva en ejecución. Los sondeos de funcionamiento y preparación permiten detectar si un servicio está en funcionamiento y listo para entregar tráfico. Ambos son tipos de verificaciones de estado y no sirven para supervisar el SLO de disponibilidad.

B. Esta respuesta es correcta. La disponibilidad es el porcentaje de tiempo que un sistema está en ejecución y puede procesar solicitudes. Cuando se supervisa esa métrica, se habilita la derivación de la métrica de SLO.

La verificación de estado no es la respuesta correcta, ya que determina si la aplicación está disponible en un momento determinado, pero no cuánto tiempo lleva en ejecución. Los sondeos de funcionamiento y preparación permiten detectar si un servicio está en funcionamiento y listo para entregar tráfico. Ambos son tipos de verificaciones de estado y no sirven para supervisar el SLO de disponibilidad.

Más recursos

Monitoring

<https://cloud.google.com/monitoring>

Administración de costos

<https://cloud.google.com/cost-management/>

Los vínculos otorgan acceso a algunos recursos útiles sobre Monitoring y la administración de costos.