

## PROYECTO No. 2 - DESPLIEGUE BÁSICO EN LA NUBE

### MIGRACIÓN DE UNA APLICACIÓN WEB A LA NUBE PÚBLICA

#### OBJETIVOS

- Conocer cómo se realiza el despliegue de aplicaciones web sobre servidores virtuales que son ejecutados en un proveedor IaaS público.
- Configurar máquinas virtuales (instancias) utilizando Compute Engine de GCP, con el fin de ejecutar una aplicación web tradicional en los servicios de la nube pública de Google.
- Configurar diferentes tipos de roles de procesamiento sobre un proveedor IaaS, específicamente los referentes a Web Server, Worker y File Server.
- Configurar y utilizar un servicio de base de datos auto-administrado por un proveedor IaaS público como es el caso del servicio Cloud SQL disponible en GCP.

#### TIEMPO DE DEDICACIÓN

La entrega “Despliegue básico en la nube” se proyecta para una semana, en la que cada estudiante deberá invertir las horas definidas en la planeación semanal del curso.

#### FECHA DE ENTREGA

La entrega deberá ser realizada de acuerdo con las fechas publicadas en la plataforma.

La socialización del proyecto estará a cargo de uno de los miembros del equipo de trabajo seleccionado aleatoriamente, por ello es importante que la totalidad de los integrantes conozcan en detalle el funcionamiento de toda la aplicación. Tenga en cuenta que, para el día de esta sustentación, **el modelo de despliegue debe encontrarse funcionando sobre GCP.**

#### LECTURAS PREVIAS

Material de lectura entregado durante la semana y documentación disponible por el proveedor GCP para los servicios de Compute Engine y Cloud SQL.

#### ESQUEMA DE EVALUACIÓN

La calificación de la entrega se distribuye de la siguiente manera:

- Actividades requeridas para la migración inicial de la aplicación: **70%**
- Documento de escenarios y resultados de las pruebas de estrés: **20%**
- Documento de la arquitectura de la aplicación: **10%**

## LUGAR Y FORMATO DE ENTREGA

La entrega está conformada por:

- Aplicación desplegada y en ejecución sobre GCP.
- Documentación de la aplicación, con el conjunto de instrucciones necesarias para su ejecución.

Por lo cual, se solicita:

- Crear un release del código fuente en el repositorio del grupo en GitHub/GitLab.
- Entregar toda la documentación vía GitHub/GitLab.

## DOCUMENTACIÓN – ARQUITECTURA DE LA APLICACIÓN

Se deberá entregar un documento donde se describa la arquitectura refinada, las conclusiones derivadas de las pruebas de estrés ejecutadas y las consideraciones adicionales para continuar escalando la aplicación web y así atender a los cientos de usuarios finales que van a subir archivos o consumir recursos del API REST de manera concurrente. En otras palabras, incluya las modificaciones, en cuanto a las características de los recursos empleados, por ejemplo, para que la siguiente versión del proyecto pueda admitir más clientes y tareas procesadas. En dicho documento se deben incluir las limitaciones del desarrollo realizado.

El nombre de este documento deberá ser: **“P2-Arquitectura de la aplicación”**.

## RECOMENDACIONES Y CONSIDERACIONES

En este proyecto, la aplicación web (API REST) desarrollada previamente será migrada a GCP. La aplicación deberá tener como mínimo cuatro componentes para su ejecución en GCP: web server, NFS server, worker y base de datos. Podrán utilizar como sistema manejador de bases de datos cualquiera de los soportados en el servicio Cloud SQL.

NO se implementarán mecanismos de autoscaling. Se recomienda desplegar las instancias exclusivamente cuando se vayan a lanzar pruebas sobre GCP y detenerlas (acción *stop*) cuando no las requieran, con la intención de evitar que se consuman la totalidad de los créditos asignados.

En el entorno de desarrollo se sugiere emplear el mismo sistema manejador de base de datos que se va a utilizar en Cloud SQL. De igual forma, el web server, worker y file server pueden ejecutarse en la máquina de desarrollo a fin de que se puedan llevar a cabo pruebas locales sin requerir del despliegue de cambios de código continuamente a GCP, lo cual puede consumir bastante tiempo.

Al finalizar y subir la entrega, elimine la base de datos de Cloud SQL, puesto que este servicio es muy costoso. En caso de solicitársele sustentación síncrona, recree la base de datos antes del encuentro con los tutores.

Dado que durante el proyecto van a tener que trabajar con credenciales de acceso a GCP, se recomienda NO copiar las credenciales de GCP en texto plano en el código fuente de la aplicación, sino emplear variables de entorno. Para su seguridad, active las alarmas de consumo y los presupuestos para monitorear de forma continua los costos asociados a su cuenta de GCP.

**Nota:** Para la sustentación de la entrega deben tener disponibles videos de prueba con distintos atributos de relación de aspecto y duración que permitan evidenciar las funcionalidades de la aplicación.

### MODELO DE DESPLIEGUE BÁSICO EN LA NUBE (APLICACIÓN EN GCP)

La compañía desea migrar la aplicación a un proveedor IaaS público como es el caso de GCP. La compañía ya ha identificado que para esta primera migración va a aprovechar los siguientes servicios:

- **Compute Engine:** Ejecución del servidor web, worker y el almacenamiento NFS. Por decisión de negocio se seleccionaron instancias de cómputo con **2 vCPU, 2 GiB en RAM y 20 GiB en almacenamiento**.
- **Cloud SQL:** Almacenamiento de los datos de la aplicación en una base de datos relacional. Seleccione una base de datos de desarrollo (Development) para minimizar los costos. En las fases iniciales de la implementación, puede usar una instancia de Compute Engine y reemplazarla por Cloud SQL solo para las pruebas de estrés.

Además, el modelo de la aplicación definido por la empresa contempla las siguientes actividades:

1. **(50%)** Se solicita el despliegue del **componente web, el worker y el sistema de almacenamiento de archivos en tres instancias de cómputo diferentes** (máquinas virtuales). Por ende, se necesita:
  - a. Instalar, configurar y ejecutar todas las dependencias (servidor de aplicaciones, librerías, etc.) y herramientas requeridas (firewall, llaves de acceso, etc.) para desplegar la aplicación web en una instancia de GCP denominada **Web Server**.
  - b. Instalar, configurar y ejecutar el componente **worker** que procesa los archivos en la segunda instancia de GCP. Para que el escenario sea comparable, establezca el mismo mecanismo y la misma tasa de procesamiento de archivos por unidad de tiempo.
  - c. Instalar, configurar y ejecutar un sistema de archivos de red (NFS) en la tercera instancia de GCP. Dicho elemento se denominará **File Server** y en él los componentes **web** y **worker** deberán almacenar y acceder a todos los archivos originales y procesados. Se deberá realizar la instalación en los componentes **web**

y **worker** del cliente NFS para que ambos puedan ingresar y guardar los archivos en el **File Server**.

**Nota:** En la capa web y worker emplee instancias con las siguientes especificaciones técnicas: **2 vCPU, 2 GiB en RAM y 20 GiB en almacenamiento**. En cuanto al NFS, puede utilizar una instancia de las mismas características o incluso una instancia micro es suficiente.

2. **(10%)** Configurar el servicio Cloud SQL para administrar la base de datos de la aplicación web. Lo anterior exige modificar el web server y el worker para que ahora todos los datos transaccionales se gestionen en la base de datos creada. No es necesario definir ningún mecanismo de replicación ni alta disponibilidad adicional.
3. **(10%)** La aplicación web debe satisfacer la totalidad de los requerimientos funcionales que ya fueron estipulados y detallados en el enunciado del proyecto.

### **ANÁLISIS DE CAPACIDAD**

Lea el documento anexo **“Entrega - Análisis de Capacidad”** para conocer toda la especificación del análisis que debe realizar a la aplicación. Los porcentajes de esta actividad se dividen en los siguientes ítems de calificación:

- **(10%)** Pruebas de estrés, análisis y documentación - Escenario 1.
- **(10%)** Pruebas de estrés, análisis y documentación - Escenario 2.

El nombre del documento a entregar deberá ser: **“P2-Análisis de capacidad”**.