**Documentation:**

Setting up Google cloud machine and start demo and API service

# 24-03-2024

****

This document provides all information to create and set up a Google Cloud Platform (GCP) virtual machine (VM) with Docker. You'll start by creating a VM instance in the Google Cloud Console, then you'll install Docker on that instance, and finally, you'll set up your Python service to run inside a Docker container. Here's a step-by-step guide:

# Step 1: Create a VM Instance in Google Cloud Platform

* Log in to Google Cloud Console: Go to the Google Cloud Console.
* Select or Create a Project: Choose an existing project or create a new one.
* Navigate to Compute Engine: In the left sidebar, navigate to "Compute Engine" > "VM instances".
* Create Instance: Click on the "Create Instance" button. You'll be taken to the instance creation form.
* Configure the instance: Choose the machine type, boot disk (e.g., an OS like Ubuntu), and allow HTTP/HTTPS traffic under the firewall settings. For running Docker and a Python service, a small to medium machine type is usually sufficient, but this depends on specific requirements. We chose the following:
  + e2-standard-4
  + 4 vCPU
  + 16 GB RAM
  + 100GB disk
* Create: After configuring your instance, click the "Create" button. Wait for a few moments until the instance is ready.

# Notes

* IP: 34.163.137.242
* Resources can be customized to drop cost
* NO-IP dynamic dns already setup, available via: anstranslation.ddns.net

# Step 2: Install Git, get the service from Git, install Python

* sudo apt install git
* git clone https://github.com/y3nk0/translation-service-ans.git
* sudo apt install python3.11-venv

# Step 3: Install Docker on the VM

After your VM instance is running, you'll need to install Docker:

* SSH into Your VM: In the VM instances dashboard, find your new VM and click the SSH button to open a terminal session directly in your browser.
* Update Your Package List (if using a Debian/Ubuntu-based OS):
  + sudo apt-get update
* Install Docker:
  + sudo apt-get install docker.io
  + If not successful follow: https://docs.docker.com/engine/install/ubuntu/#install-using-the-repository
* Verify Docker Installation:
  + sudo docker --version: This command should return the Docker version, indicating that Docker has been successfully installed.

# Notes

Model checkpoints need to be stored in models directory

# Step 4: Run the service

* Go the demo\_n\_api directory: cd demo\_n\_api/
* Build Your Docker Image: sudo docker compose up --build
* To shut down service: sudo docker compose down -v
* Service will be available at: http://34.163.137.242 and http://anstranslation2.ddns.net (without using https)