



Computação na Nuvem

Licenciatura em Engenharia
Informática, Redes e Telecomunicações

Relatório do Trabalho Final

Grupo 06

Docentes:

Prof. José Simão
Prof. Luís Assunção

Discentes:

Francisco Chicharro nº43861
João Florentino nº43874

Semestre de Verão 19/20

21 junho de 2020

Conteúdo

Lista de Figuras	3
Objetivos	4
Descrição	5
Serviços da Google Cloud Platform	6
Tipo de serviço e elasticidade	6
Fase 1: Comunicação entre utilizador e servidor	7
Fase 2: Comunicação entre servidor e OCR	9
Fase 3: Tradução do texto – Comunicação OCR/Translate	11
Métrica utilizada para escalonamento das máquinas virtuais (VM's)	13
Problemas encontrados e resolução	13
Conclusão	14

Lista de Figuras

Figura 1: Sistema CNText e respectivas interações	5
Figura 2: Comunicação cliente-Servidor	7
Figura 3:Módulo servidor	8
Figura 4:Comunicação pub/sub	9
Figura 5:Acesso aos serviços GCP através das VMs	10
Figura 6:Mensagens Pub/Sub entre TextOCR e Translate	11
Figura 7: Comunicação com Translate	12

Objetivos

Este trabalho tem como objetivo planejar e realizar um sistema, designado por CNText, que permita detetar o texto presente numa imagem, sendo posteriormente possível traduzir esse mesmo texto para diferentes linguagens. O sistema tem a capacidade de submeter e executar tarefas, possuindo requisitos de elasticidade que permitem aumentar e diminuir a capacidade de processamento de imagens, utilizando as funcionalidades dos serviços da Google Cloud Platform de armazenamento e comunicação, nomeadamente os serviços: Storage, Firestore e Pub/Sub e Compute Engine.

Descrição

O sistema CNText é composto por várias fases que são implementadas de forma isolada, sendo que cada componente terá uma função específica no funcionamento do sistema:

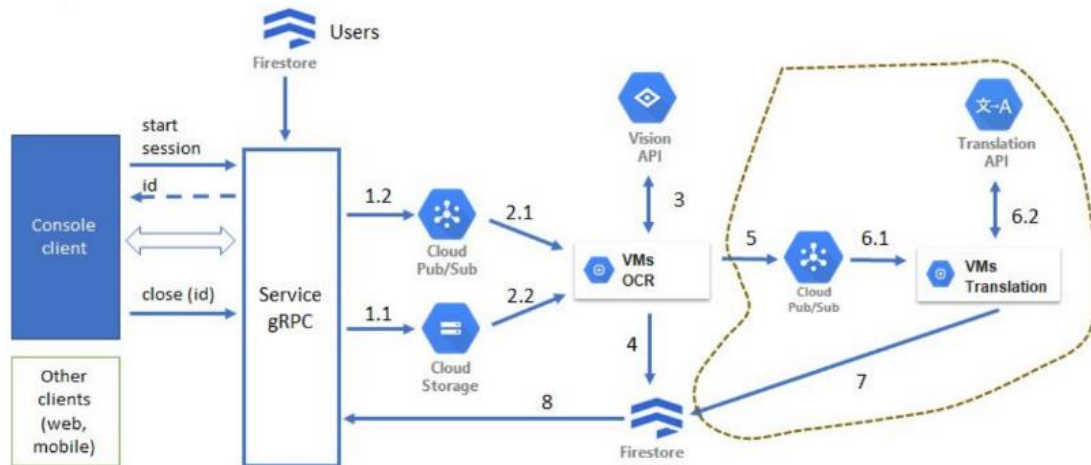


Figura 1: Sistema CNText e respectivas interações

- Numa primeira fase, começa-se por tratar a autenticação dos utilizadores no sistema, onde estes iniciam uma sessão através do cliente gRPC. Os utilizadores estão previamente criados na coleção "Users" no serviço de Firestore onde estão identificados pelo seu nome de utilizador e o nível de serviço (Free ou Premium). O serviço gRPC vai retornar o identificador da sessão iniciada sendo este guardado pelo utilizador.
- Após autenticação, é iniciado o processo de submissão de uma imagem (1.1), esta é guardada no serviço Storage sendo retornado ao cliente gRPC um identificador único para que seja possível pedir a tradução do texto da imagem.
- Posteriormente, os pedidos de deteção de texto (OCR) chegam às VMs através do serviço de Pub/Sub (1.2) onde existem dois tópicos: free-ocr e premium-ocr. Ao tópico free-ocr estão associados os pedidos free, onde uma subscrição é consumida por uma única VM enquanto que ao tópico premium-ocr estão associados os pedidos premium, onde uma subscrição é consumida por várias VMs.
- De seguida, a cada VM de OCR vai receber o nome da imagem que tem de processar (2.1) onde irá obter o conteúdo desta a partir do Storage (onde foi anteriormente guardada) (2.2). Posteriormente, este processo irá interagir com a API de Visão (3) que deteta o texto de uma imagem. Este texto após ser detetado, será guardado numa coleção do Firestore (4).

- Após este processo, a imagem é apagada do Storage e será enviada uma notificação com o resultado de detecção OCR, para um tópico Pub/Sub (5).

O trabalho está dividido em 3 módulos distintos:

- ➔ TextClient: Onde se implementam as operações que correspondem ao Cliente.
- ➔ TextServer: Onde se implementam as operações que correspondem ao Servidor (Sistema)
- ➔ TextOCR: Módulo que trata o serviço de detecção de texto assim como todas as ligações necessárias para o estabelecimento do serviço.

Serviços da Google Cloud Platform

- **Storage:** Serviço responsável pelo armazenamento de imagens a processar num bucket criado previamente e especificado para o cliente. As imagens serão removidas quando não forem necessárias.
- **Pub/Sub:** Serviço de troca de mensagens que servirá de canal de comunicação entre o Servidor e as máquinas virtuais (VMs) de OCR.
- **Firestore:** Serviço de armazenamento de dados correspondentes ao utilizador podendo ser consultados, sendo possível guardar documentos com informações referentes às imagens submetidas pelos utilizadores, como o texto detetado, a linguagem desse texto detetada.
- **Compute Engine:** Serviço que permite a criação de máquinas virtuais (VMs) e os grupos de instância que executam o sistema para detecção de texto. Permitem a utilização da função de escalonamento de grupos de instância.

Tipo de serviço e elasticidade

O sistema CNText oferece dois níveis de qualidade de serviço: free e premium onde os utilizadores estão associados a um dos níveis de serviço. Como dito anteriormente, este tipo de associação é estabelecido previamente na coleção de users no serviço Firestore.

Os níveis de serviço vão influenciar a forma como os pedidos de tradução de cada utilizador são processados:

- Para tratar os pedidos do nível free existe apenas 1 VM de OCR e 1VM de tradução, assim, só é iniciado o processamento de um pedido após conclusão do anterior
- Para tratar os pedidos do nível premium existe um número variável de VMs (de instance groups) para OCR e tradução, isto é , pedidos em simultâneo

Fase 1: Comunicação entre utilizador e servidor

A aplicação cliente TextClient irá conter todos os métodos necessários para a interação com o utilizador: Início e fim de sessão (login() e logout()), submissão de imagem (sendImage()), deteção de texto (detectText()) e tradução do texto da imagem.

A aplicação servidor TextServer será responsável pela interação com os serviços da Google Cloud Platform, sendo necessário a utilização do serviço Storage para armazenar a imagem, o serviço Firestore para aceder á coleção de documentos (registando a autenticação dos utilizadores e apresentando o processamento de imagens) e por fim, pelo escalonamento dos instance groups no tipo de serviço premium.

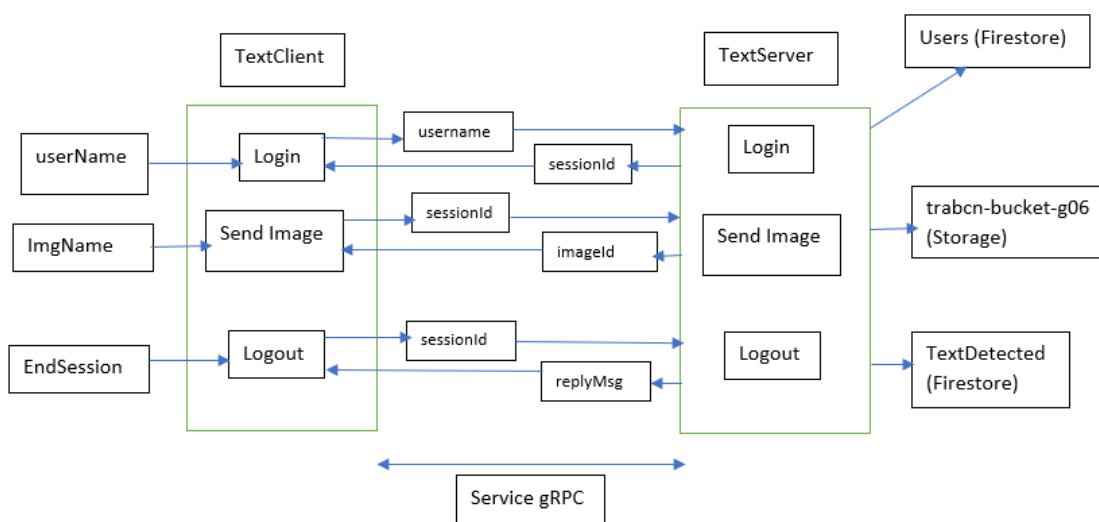


Figura 2: Comunicação cliente-Servidor

É apresentado então um Menu na consola da aplicação TextClient com as seguintes opções (Figura 2):

Login: Para um utilizador iniciar sessão, necessita de apresentar na consola o seu username. Se o seu username estiver registado no sistema (Coleção “Users”) então o cliente efetua o login com sucesso. A este cliente, é passado um ID de sessão para que o servidor tenha conhecimento deste cliente, sendo possível a distinção de múltiplos clientes com sessão iniciada. Após o login com sucesso, é apresentado ao utilizador um menu de opções disponíveis no sistema

Submissão de imagem: Para submeter uma imagem, o cliente terá de anunciar o nome da imagem que quer submeter. É passado então na consola o caminho da pasta onde a imagem se localiza seguido do nome desta. De seguida, é pedido ao utilizador a linguagem que se pretende traduzir o texto. A aplicação cliente atribui assim um identificador á imagem que se quer submeter distinguindo as imagens que este utilizador possa submeter. Assim, do lado servidor, assim que este possuir a imagem irá criar um identificador no formato de “Imagem-<ID de sessão>_<ID de

imagem>_<Linguagem> (Exemplo: Imagem-1_1_en) enviando para a aplicação cliente o formato de como a imagem foi registrada.

Terminar sessão: Quando um utilizador deseja terminar a sessão é enviado para o servidor o respetivo identificador de sessão para avisar o servidor para que este atualize os registos das sessões ativas. Posto isto, a aplicação cliente encerra.

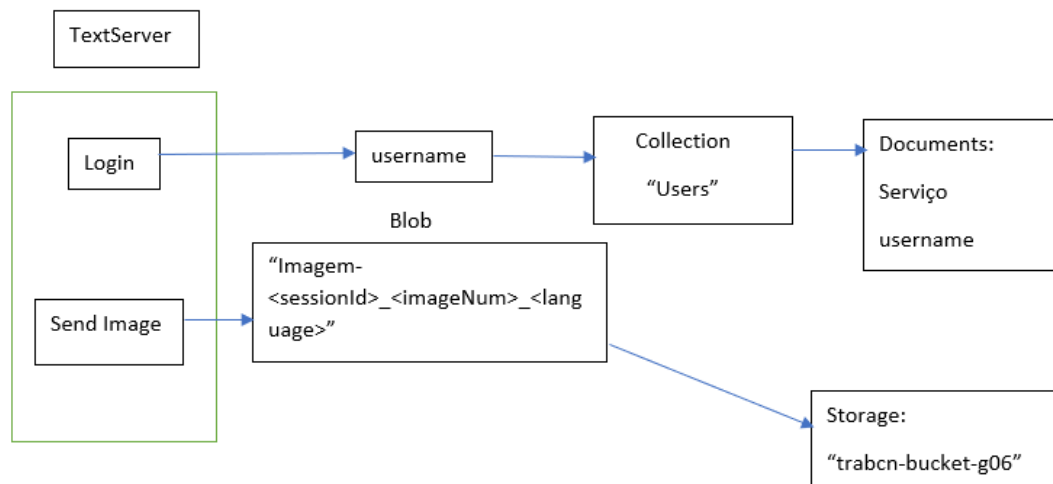


Figura 3: Módulo servidor

De seguida, analisa-se o comportamento entre o Servidor e o acesso aos serviços presentes na Google Cloud Platform (Figura 3):

Login: Quando o utilizador submete o seu username, o servidor irá aceder á coleção de "Users" do serviço Firestore para validar o utilizador. Caso este não se encontre na lista dos documentos, a aplicação cliente é notificada sobre a inexistência do mesmo.

Submissão de Imagem: No processo da submissão de imagem no serviço Storage, criou-se um bucket geral "trabcn-bucket-g06" e, para cada imagem submetida, é criado um blob que tem exatamente o mesmo nome da imagem, facilitando assim o processo da colocação de imagem no Storage.

Fase 2: Comunicação entre servidor e OCR

Após a submissão de um ficheiro correspondente a uma imagem pelo cliente, este é enviado para uma VM OCR. O processo é feito com ajuda do serviço Pub/Sub da GCP

O servidor para guardar registos de clientes com a sessão iniciada, distingue os utilizadores free e premium. Assim, quando se submeter um pedido para a tradução da imagem, o servidor irá publicar ou no tópico “free-ocr” ou no “premium-ocr” onde irá colocar num destes a mensagem que contém o nome do Blob criado a partir da submissão de imagem e a linguagem para qual se pretende traduzir. O subscritor free “free-ocr-message” e “premium-ocr-message” aguardam pelas publicações nos tópicos (Figura 4):

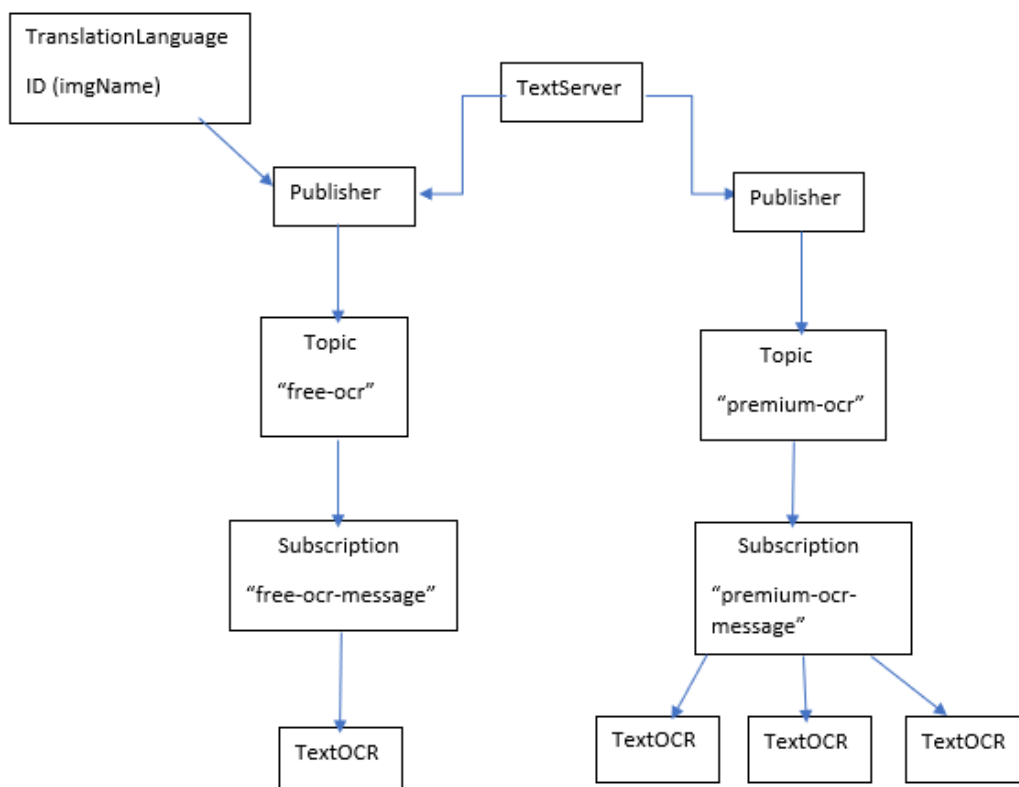


Figura 4: Comunicação pub/sub

Assim, seguem-se os passos necessários para a deteção de texto na imagem (Figura 5):

Download do Blob criado: Os subscritores free e premium recebem o nome do Blob para fazer o download do mesmo. Assim, o subscritor acede ao bucket “trabcn-bucket-g06”, onde serão retornados “bytes” do ficheiro que serão guardados numa pasta definida (ImagesOCR)

Deteção de texto na imagem: Uma vez tendo a imagem, a VM irá enviá-la para a VisionAPI para ser possível extrair o texto que ela possui. Após o envio, a VM de OCR irá obter o texto detetado.

Criação do documento: Após deteção de texto, a VM vai aceder á coleção “Text-Detected” onde irá criar um documento que tem o mesmo nome do Blob e com os campos de linguagem detetada, texto detetado e texto traduzido. O campo de translatedText será apenas preenchido quando se realizar a tradução no passo seguinte:

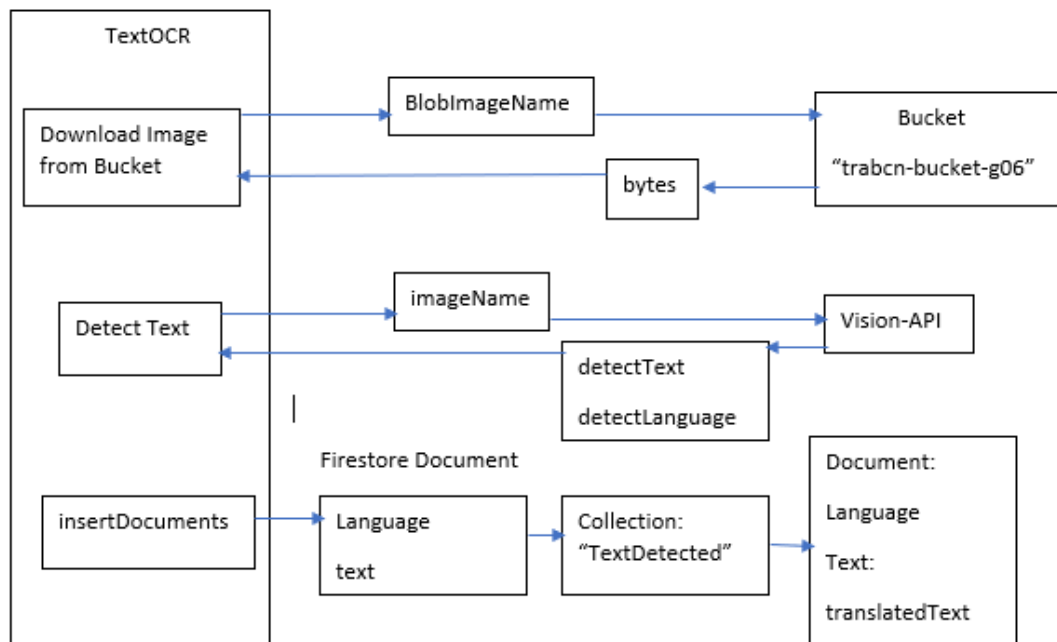


Figura 5: Acesso aos serviços GCP através das VMs

Fase 3: Tradução do texto – Comunicação OCR/Translate

Assim que a VM de OCR terminar o processo de detecção de texto, irá enviar um publish para o tópico “free-translate” ou para “premium-translate” com o nome do documento criado que contém o nome da imagem, linguagem detetada e para qual se quer traduzir. As subscrições correspondentes aos tópicos mencionados vão processar a mensagem (Figura 6):

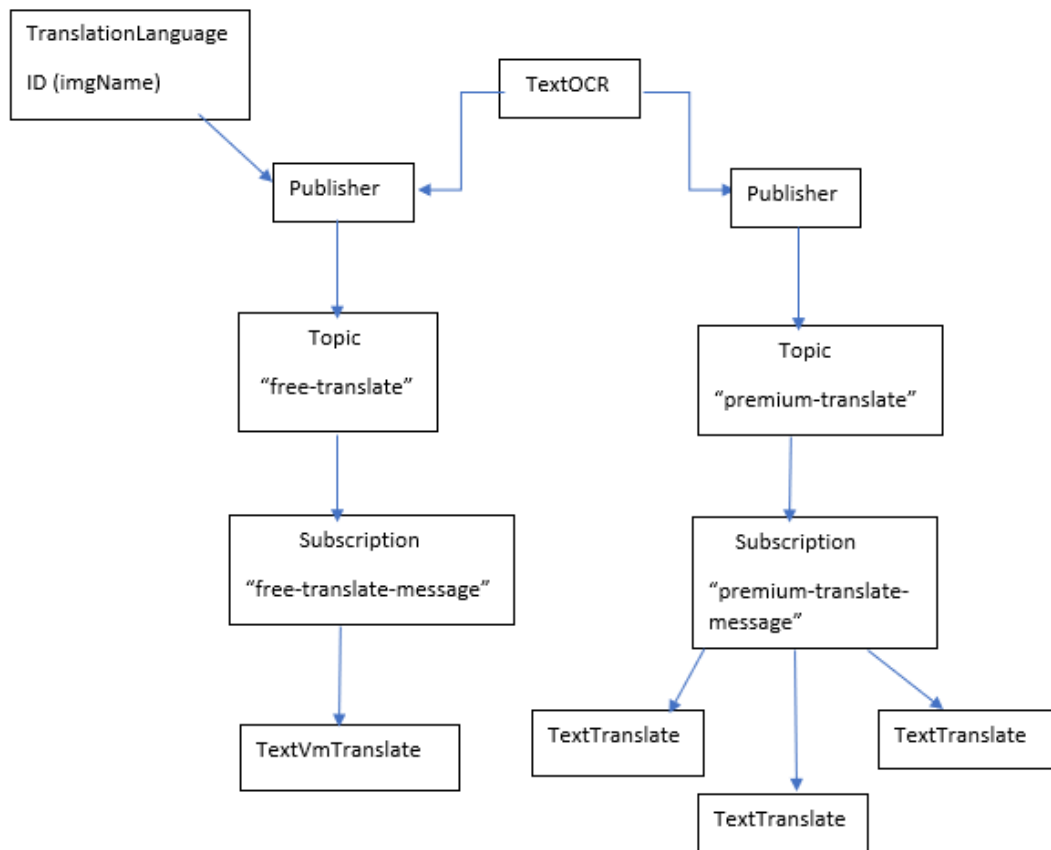


Figura 6: Mensagens Pub/Sub entre TextOCR e Translate

A VM de tradução terá o seguinte comportamento (Figura 7):

Receber o documento: A VM recebe a mensagem do serviço Pub/Sub obtendo o nome do documento da coleção “TextDetected” tendo assim acesso ao texto detetado

Pedido de tradução: Utilizando os campos da mensagem e o texto, é feito um pedido à aplicação Translation API, recebendo o texto traduzido.

Alteração do documento: Tendo agora o texto traduzido, a VM translate irá alterar o campo “translatedText” para este possuir agora o texto traduzido (encontrava-se vazio anteriormente)

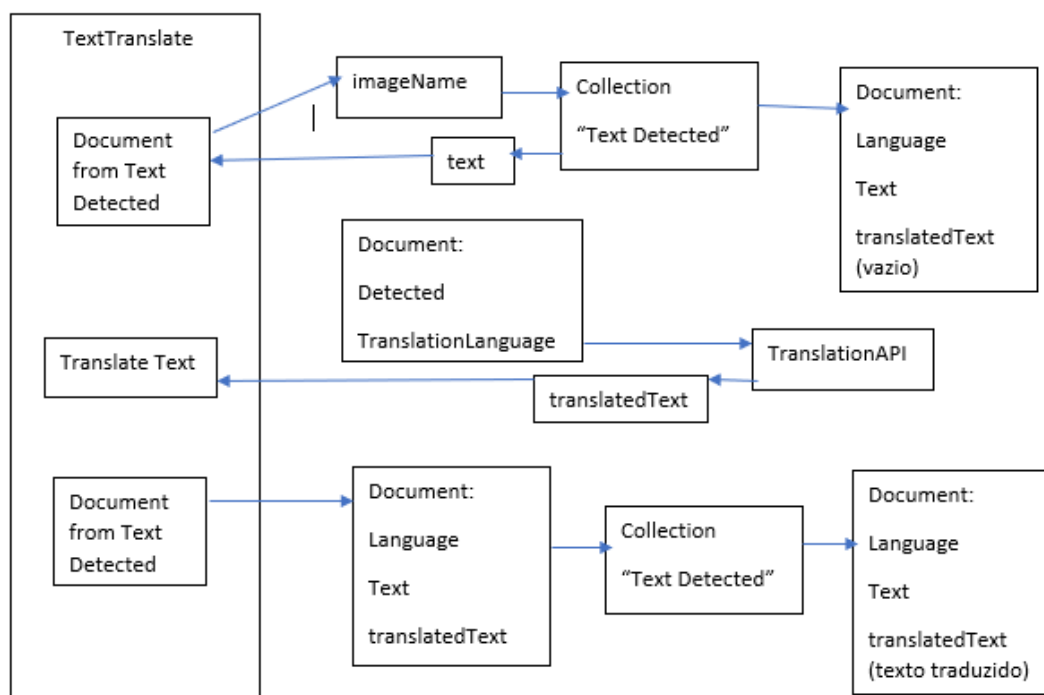


Figura 7: Comunicação com Translate

Métrica utilizada para escalonamento das máquinas virtuais (VM's)

Nível de serviço: Free

As máquinas utilizadas em VMs free possuem uma máquina do tipo f1 -micro com 1 CPU, memória de 0.6Gb. Esta VM possui apenas uma thread.

Nível de serviço: Premium

As máquinas utilizadas em VMs premium são do tipo e2-micro com 2 CPUs, memória de 1Gb. Cada subscritor deste tipo de VM usa duas threads. Este nível de serviço faz escalonamento de instance groups (grupos de instância) segundo a regra:

$N^{\circ}Vms = n^{\circ}. Utilizadores$

Assim, para efeitos de teste, considerou-se o aumento de uma instância por cada início de sessão de um cliente premium. Terminando a sessão da mesma, verificou-se que a instância é retirada no instance-groups.

Problemas encontrados e resolução

Quando se colocou o sistema a correr sobre as máquinas virtuais (Free e Premium) até ao momento as credenciais da google (GOOGLE_CREDENTIALS) estava declaradas no código passando o caminho completo, tendo assim apenas acesso local. Foi necessário alterar a forma como se iniciava as credenciais da conta google através do método `getApplicationDefault()`.

A parte da tradução do texto detetado foi testado inicialmente apenas localmente uma vez que a VM translate tinha uma versão não atualizada no sistema, sendo que o erro foi detetado mais tarde e corrigido sendo possível a utilização desta para tradução de texto.

Conclusão

Com este trabalho conseguiu-se aprofundar os conhecimentos práticos e teóricos que têm vindo a ser apresentados ao longo do semestre, usufruindo da utilização de vários serviços da Google Cloud Platform para realizar um sistema de deteção de texto de uma imagem assim como traduzi-lo para uma linguagem á escolha do utilizador nomeadamente os serviços de Storage, Firestore, PubSub e Compute Engine para Instanciação de máquinas virtuais.

Assim, foi possível criar um sistema onde um utilizador realiza início de sessão onde lhe é apresentado um menu com várias opções de ações a realizar: Envio de uma imagem, deteção de texto da imagem e de seguida a tradução da mesma. Todo este processo necessita da confirmação nos vários serviços utilizados de modo a que o pedido é feito de forma ordeira e correta, chegando então ao objetivo final : A obtenção do texto traduzido