

Infraestructura para Big Data

Práctica 6 - IBM Cloud

José Manuel Bustos Muñoz

Realice las siguientes acciones documentando breve y concisamente el proceso:

1. **Active un bucket para almacenamiento de objetos en el cloud (COS). Genere credenciales para permitir el acceso mediante servicios externos que sean compatibles con Amazon S3.**

Accedemos a IBM Cloud, y buscamos en el catálogo la categoría de Almacenamiento, accediendo a ella creamos el servicio de almacenamiento de objetos en el cloud (COS).

IBM Cloud

Buscar recursos y ofertas...

Q

Catálogo

Documentos

Soporte

Gestionar

Jose Manu

Catálogo

Q label:lite

Todas las categorías (57)

Cálculo (10)

Contenedores (1)

Gestión de redes

Almacenamiento (1) >

AI (15)

Análisis (4)

Bases de datos (3)

Herramientas de desarrollador (7)


Integración (3)

Internet de las cosas (1)

Seguridad e identidad (2)

Visualización (6)

Almacenamiento

 **Almacenamiento de objetos**
Lite • IBM • Habilitado para IAM

Proporciona flexibilidad, rentabilidad y almacenamiento escalable en la nube para datos no estructurados.

IBM Cloud

Curso: Infraestructura para Big Data (2017-19)

Buscar recursos y ofertas...

Q

Catálogo

Documentos

Soporte

Gestionar

Jose Manuel Bustos Muñoz'...

Ver todo

 **Almacenamiento de objetos en la nube**
Lite • IBM

IBM Cloud Object Storage es un servicio de almacenamiento en la nube altamente escalable, diseñado para una durabilidad, resiliencia y seguridad de primer nivel. Almacene, gestione y acceda a sus datos a través de nuestro portal autoservicio y de las API RESTful. Conecte aplicaciones directamente con Cloud Object Storage y utilice otro servicio de IBM Cloud con sus datos.

Ver documentos

Condiciones

AUTOR

IBM

PUBLICADO

09/05/2019

TIPO

Servicio

Nombre del servicio:

Almacenamiento de objetos en la nube -bj

Seleccione un grupo de recursos: ⓘ

Default

Etiquetas: ⓘ

Ejemplos: env:dev, version-1

Características

• Almacenamiento para IBM Cloud

IBM Cloud Object Storage proporciona almacenamiento de datos no estructurados para aplicaciones en la nube. Las bibliotecas y los SDK ofrecen soporte a un conjunto común de funciones de API compatibles con S3 para

• Transferencia de alta velocidad de Aspera integrada

Con la transferencia de datos de alta velocidad de IBM Cloud Object Storage Aspera, puede mejorar el rendimiento de la transferencia de datos al transferir rápidamente los datos de largas distancias y bajo diversas condiciones de red. Se

¿Necesita ayuda?

Póngase en contacto con el soporte de IBM Cloud ⓘ

Añadir a estimación

Crear

Una vez creado el servicio COS ahora generamos el depósito, y luego las credenciales para utilizar el servicio.

IBM Cloud

Buscar recursos y ofertas...

Catálogo Documentos Soporte

Getting started

Buckets

Endpoint

Credenciales de servicio

Conexiones

Usage details

Plan

Lista de recursos /

Almacenamiento de objetos en la nube -bj

Grupo de recursos: Default [Añadir etiquetas](#)

Crear depósito

Nombre de depósito exclusivo: [Consultar reglas de denominación](#)

depositoejemplo

Resiliencia	Ubicación
Regional	eu-gb

Mejor rendimiento

Clase de almacenamiento [Consulte los precios para cada clase](#)

Estándar

☐ Añadir claves de protección de clave (opcional)

☐ Añadir política de archivado (opcional)

☐ Añadir política de retención (opcional)

Esta característica solo está disponible para nuestros clientes del plan Estándar. [Ver precios](#)

Añadir nueva credencial

Nombre:

Credenciales de servicio-1

Rol:

Writer

Seleccione el ID de servicio (opcional)

Generar automáticamente

☒ Include HMAC Credential

Añadir parámetros de configuración en línea (opcional):

{"HMAC":true}

Cancelar

Añadir

Se pueden visualizar las credenciales:

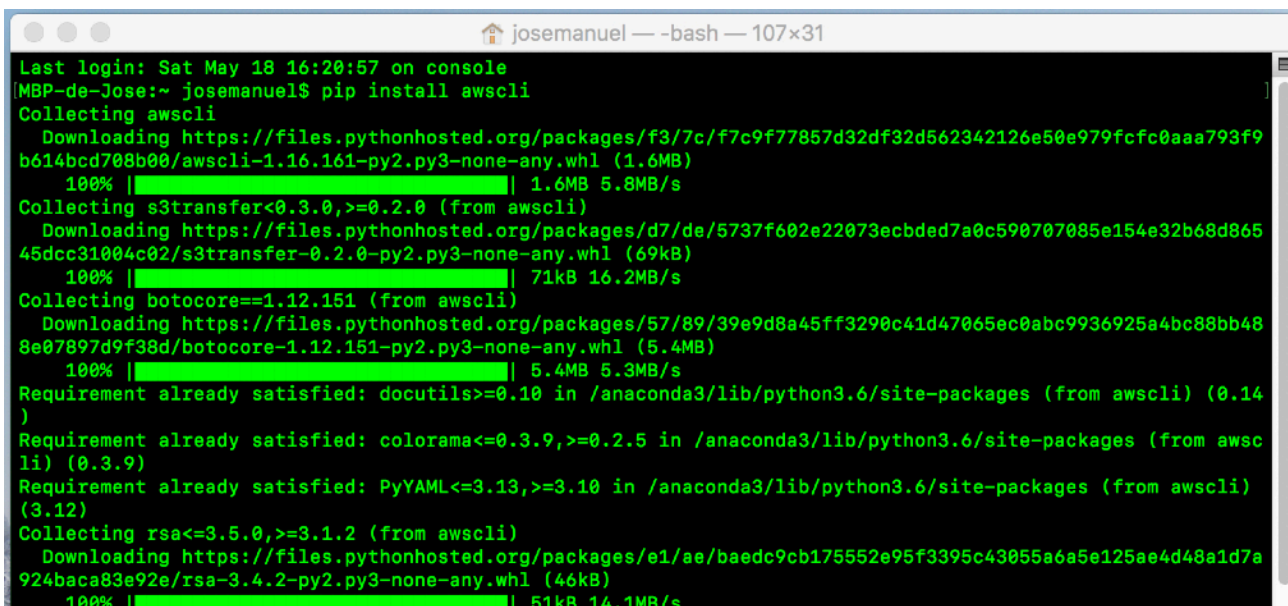


The screenshot shows the IBM Cloud Object Storage console. At the top, it says "Almacenamiento de objetos en la nube -bj". Below that, there's a "Grupo de recursos: Default" and a button "Añadir etiquetas". On the right, there's a link "Ver documentos". The main section is titled "Credenciales de servicio" and has a "Nueva credencial" button. Below this, there's a table with columns: "NOMBRE DE CLAVE", "FECHA DE CREACIÓN", and "ACCIONES". The table has one row: "Credenciales de servicio-1", "18 DE MAY. DE 2019 - 05:01:28 PM", and "Ver credenciales". Below the table, there's a JSON object showing the credentials:

```
{
  "apikey": "EfUvQ56_7hg9uUGkWB9Etf4-rSD7lG_aK69qxsq0MaMS",
  "cos_hmac_keys": {
    "access_key_id": "d8154d13656341b597ba0dade81d243c",
    "secret_access_key": "91fa34e926c07056acf72407cbcb6bd9c6de44ed312afc4"
  },
  "endpoints": "https://control.cloud-object-storage.cloud.ibm.com/v2/endpoints",
  "iam apikey description": "Auto-generated for key d8154d13-6563-41b5-97ba-0dade81d243c".
}
```

2. Utilizando la CLI de AWS, cargue un fichero CSV tabular en el bucket creado anteriormente.

Instalamos el "awscli" en el pc local:



```
josemanuel — bash — 107x31
Last login: Sat May 18 16:20:57 on console
MBP-de-Jose:~ josemanuel$ pip install awscli
Collecting awscli
  Downloading https://files.pythonhosted.org/packages/f3/7c/f7c9f77857d32df32d562342126e50e979fcfc0aaa793f9b614bcd708b00/awscli-1.16.161-py2.py3-none-any.whl (1.6MB)
    100% |#####| 1.6MB 5.8MB/s
Collecting s3transfer<0.3.0,>=0.2.0 (from awscli)
  Downloading https://files.pythonhosted.org/packages/d7/de/5737f602e22073ecbde7a0c590707085e154e32b68d86545dcc31004c02/s3transfer-0.2.0-py2.py3-none-any.whl (69kB)
    100% |#####| 71kB 16.2MB/s
Collecting botocore==1.12.151 (from awscli)
  Downloading https://files.pythonhosted.org/packages/57/89/39e9d8a45ff3290c41d47065ec0abc9936925a4bc88bb488e07897d9f38d/botocore-1.12.151-py2.py3-none-any.whl (5.4MB)
    100% |#####| 5.4MB 5.3MB/s
Requirement already satisfied: docutils>=0.10 in /anaconda3/lib/python3.6/site-packages (from awscli) (0.14)
Requirement already satisfied: colorama<=0.3.9,>=0.2.5 in /anaconda3/lib/python3.6/site-packages (from awscli) (0.3.9)
Requirement already satisfied: PyYAML<=3.13,>=3.10 in /anaconda3/lib/python3.6/site-packages (from awscli) (3.12)
Collecting rsa<=3.5.0,>=3.1.2 (from awscli)
  Downloading https://files.pythonhosted.org/packages/e1/ae/baedc9cb175552e95f3395c43055a6a5e125ae4d48a1d7a924baca83e92e/rsa-3.4.2-py2.py3-none-any.whl (46kB)
    100% |#####| 51kB 14.1MB/s
```

Luego con el comando “aws configure” introducimos las credenciales para conectar con el servicio COS:

```
MBP-de-Jose:~ josemanuel$ aws configure
AWS Access Key ID [None]: d8154d13656341b597ba0dade81d243c
AWS Secret Access Key [None]: 91fa34e926c07056acf72407cbcba6bd9c6de44ed312afc4
Default region name [None]: eu-gb
Default output format [None]:
```

Usando de nuevo aws cli, conectamos con COS y subimos un fichero .csv, luego podemos hacer un ‘ls’ y ver que se ha subido:

```
MBP-de-Jose:p6 infraestructura -19 mayo- josemanuel$ aws --endpoint-url=https://s3.eu-gb.objectstorage.softlayer.net s3 cp Players.csv s3://depositoejemplo
upload: ./Players.csv to s3://depositoejemplo/Players.csv
MBP-de-Jose:p6 infraestructura -19 mayo- josemanuel$ aws --endpoint-url=https://s3.eu-gb.objectstorage.softlayer.net s3 ls s3://depositoejemplo
2019-05-18 17:28:53      281111 Players.csv
MBP-de-Jose:p6 infraestructura -19 mayo- josemanuel$
```

Si accedemos a la interfaz de IBM Cloud podemos ver como en el depósito ya aparece el archivo .csv que hemos subido:

Getting started

Buckets

Objects

Configuration

Access Policies

Endpoint

Credenciales de servicio

Conexiones

Usage details

Plan

Lista de recursos / Almacenamiento de objetos en la nube -bj / depositoejemplo

Ver documento:

depositoejemplo

Objetos

Subir

<input type="checkbox"/>	Nombre de objeto	Archivado	Tamaño	Última modificación	
<input type="checkbox"/>	Players.csv		274,5 KB	18/05/2019 5:28:53 P. M.	...

Elementos por página: 10 | 1-10 of items

Arrastre y suelte archivos o carpetas para subirlos.

A partir del almacenamiento y datos cargados en el Ejercicio 1, realice las siguientes acciones documentando breve y concisamente el proceso:

3. Active el servicio de análisis de mediante SQL de datos en el COS. Realice una consulta de prueba que contabilice el número de filas en el fichero CSV previamente cargado.

Se da de alta un servicio de SQL Query, para poder realizar queries contra el fichero cargado en COS.

The screenshot shows the IBM Cloud console interface. At the top, there's a navigation bar with 'IBM Cloud' and a search bar. Below it, the 'Gestionar' (Manage) section is active, showing a list of resources. The selected resource is 'SQL Query-p8'. The details page shows the service's location (Dallas) and a button to 'Launch SQL Query UI'. The 'Deployment details' section displays the CRN and the API Endpoint. The 'Encryption' section shows the encryption method as 'Service-managed key encryption'.

Detalles de servicio - IBM Cloud

Grupo de recursos: Default Ubicación: Dallas [Añadir etiquetas](#)

Overview Docs [Launch SQL Query UI](#)

Deployment details

CRN: crn:v1:bluemix:public:sql-query:us-south:a/bb278d18ce404cec9f29545471091daa:8260150b-1394-400b-bb9f-6a92264f29bd

API Endpoint: <https://api.sql-query.cloud.ibm.com>

Encryption

Encryption method: Service-managed key encryption

Una vez generado el servicio, lanzamos el launch de SQL Query, y hacemos por ejemplo una query inicial para mostrar los datos del fichero subido a COS:

The screenshot shows the IBM SQL Query console. At the top, there's a navigation bar with 'IBM SQL Query', 'Support', 'API', 'Docs', and 'Samples'. The main area shows a query editor with a single query: 'select * from cos://eu-gb/depositoejemplo'. Below the editor, the 'Target location' is set to 'cos://us-south/sql-08dc1f9c-5926-493b-84ef-63b5a5a74c71/result/ (Initial target)'. A 'Run' button is visible. The results section shows a table with 8 columns and 3922 rows. The columns are: _CO, PLAYER, HEIGHT, WEIGHT, COLLAGE, BORN, BIRTH_CITY, and BIRTH_STATE. The rows show data for various players, including Curly Armstrong, Cliff Barker, Leo Barnhorst, Ed Bartels, Ralph Beard, Gene Berce, Charlie Black, and Nelson Bobb.

IBM SQL Query Support API Docs Samples

1 select * from cos://eu-gb/depositoejemplo

Detalles de servicio - IBM Cloud

Target location: cos://us-south/sql-08dc1f9c-5926-493b-84ef-63b5a5a74c71/result/ (Initial target) [Run](#)

Jobs	Result	Query details																																																																								
Completed Today 5:52 PM	8 columns 3922 rows (Previewing 315 rows)																																																																									
	<table><thead><tr><th>_CO</th><th>PLAYER</th><th>HEIGHT</th><th>WEIGHT</th><th>COLLAGE</th><th>BORN</th><th>BIRTH_CITY</th><th>BIRTH_STATE</th></tr></thead><tbody><tr><td>0</td><td>Curly Armstrong</td><td>180</td><td>77</td><td>Indiana University</td><td>1918</td><td></td><td></td></tr><tr><td>1</td><td>Cliff Barker</td><td>188</td><td>83</td><td>University of Kentuck</td><td>1921</td><td>Yorktown</td><td>Indiana</td></tr><tr><td>2</td><td>Leo Barnhorst</td><td>193</td><td>86</td><td>University of Notre D</td><td>1924</td><td></td><td></td></tr><tr><td>3</td><td>Ed Bartels</td><td>196</td><td>88</td><td>North Carolina State</td><td>1925</td><td></td><td></td></tr><tr><td>4</td><td>Ralph Beard</td><td>178</td><td>79</td><td>University of Kentuck</td><td>1927</td><td>Hardinsburg</td><td>Kentucky</td></tr><tr><td>5</td><td>Gene Berce</td><td>180</td><td>79</td><td>Marquette University</td><td>1926</td><td></td><td></td></tr><tr><td>6</td><td>Charlie Black</td><td>196</td><td>90</td><td>University of Kansas</td><td>1921</td><td>Arco</td><td>Idaho</td></tr><tr><td>7</td><td>Nelson Bobb</td><td>183</td><td>77</td><td>Temple University</td><td>1924</td><td>Philadelphia</td><td>Pennsylvania</td></tr></tbody></table>	_CO	PLAYER	HEIGHT	WEIGHT	COLLAGE	BORN	BIRTH_CITY	BIRTH_STATE	0	Curly Armstrong	180	77	Indiana University	1918			1	Cliff Barker	188	83	University of Kentuck	1921	Yorktown	Indiana	2	Leo Barnhorst	193	86	University of Notre D	1924			3	Ed Bartels	196	88	North Carolina State	1925			4	Ralph Beard	178	79	University of Kentuck	1927	Hardinsburg	Kentucky	5	Gene Berce	180	79	Marquette University	1926			6	Charlie Black	196	90	University of Kansas	1921	Arco	Idaho	7	Nelson Bobb	183	77	Temple University	1924	Philadelphia	Pennsylvania	
_CO	PLAYER	HEIGHT	WEIGHT	COLLAGE	BORN	BIRTH_CITY	BIRTH_STATE																																																																			
0	Curly Armstrong	180	77	Indiana University	1918																																																																					
1	Cliff Barker	188	83	University of Kentuck	1921	Yorktown	Indiana																																																																			
2	Leo Barnhorst	193	86	University of Notre D	1924																																																																					
3	Ed Bartels	196	88	North Carolina State	1925																																																																					
4	Ralph Beard	178	79	University of Kentuck	1927	Hardinsburg	Kentucky																																																																			
5	Gene Berce	180	79	Marquette University	1926																																																																					
6	Charlie Black	196	90	University of Kansas	1921	Arco	Idaho																																																																			
7	Nelson Bobb	183	77	Temple University	1924	Philadelphia	Pennsylvania																																																																			

Vemos el número de filas o registros del fichero:

IBM SQL Query

Support ▾APIDocs

1 select count(*) from cos://eu-gb/depositoejemplo

Target location: cos://us-south/sql-08dc1f9c-5926-493b-84ef-63b5a5a74c71/result/ (Initial target)

Jobs	Result	Query details
Completed Today 5:53 PM	1 column 1 row	
	COUNT(1)	
Completed Today 5:52 PM	3922	

4. Realice una segunda consulta de su elección que permita resumir o visualizar los datos en el CSV cargado. Identifique dónde y cómo se han almacenado los resultados de las consultas.

Agrupamos por universidad, y vemos el número de jugadores con altura de 2 metros o más por cada una de las universidades:

IBM SQL Query

Support API Docs

Samples

```
1 select count(*) as numHeight, collage from cos://eu-gb/depositoejemplo where height >= 200 group by collage order by numHeight desc
```

Target location: cos://us-south/sql-08dc1f9c-5926-493b-84ef-63b5a5a74c71/result/ (Initial target)

Run

Jobs	Result	Query details																		
Completed Today 6:04 PM	2 columns 311 rows																			
	<table><thead><tr><th>NUMHEIGHT</th><th>COLLAGE</th></tr></thead><tbody><tr><td>252</td><td></td></tr><tr><td>44</td><td>University of California, Los Angeles</td></tr><tr><td>36</td><td>University of Kentucky</td></tr><tr><td>33</td><td>Syracuse University</td></tr><tr><td>33</td><td>University of Kansas</td></tr><tr><td>31</td><td>Duke University</td></tr><tr><td>31</td><td>University of North Carolina</td></tr><tr><td>24</td><td>Indiana University</td></tr></tbody></table>	NUMHEIGHT	COLLAGE	252		44	University of California, Los Angeles	36	University of Kentucky	33	Syracuse University	33	University of Kansas	31	Duke University	31	University of North Carolina	24	Indiana University	
NUMHEIGHT	COLLAGE																			
252																				
44	University of California, Los Angeles																			
36	University of Kentucky																			
33	Syracuse University																			
33	University of Kansas																			
31	Duke University																			
31	University of North Carolina																			
24	Indiana University																			
Completed Today 6:03 PM																				
Completed Today 6:03 PM																				
Completed Today 6:01 PM																				
Completed Today 5:53 PM																				

Si vamos al bucket creado, además del depósito creado aparece otro de SQL.

Depósitos

Prefijo filtrado

Crear depósito +

Nombre	Acceso público	Ubicación	Clase de almacenamiento	Creado	Avanzado
depositoejemplo		eu-gb	Estándar	18/05/2019 17:00:13	...
sql-08dc1f9c-5926-493b-84ef-63b5a5a74c71		us-south	Estándar	18/05/2019 17:51:15	...

Elementos por página: 10 | 1-10 of items

< >

Accediendo a él, se pueden ver los objetos generados por las consultas anteriores, los resultados de los jobs lanzados.

Se genera un CSV por cada job, y puede descargarse y verse los datos de la consulta en cuestión.

sql-08dc1f9c-5926-493b-84ef-63b5a5a74c71

Ver documentos

Objetos

Prefijo filtrado

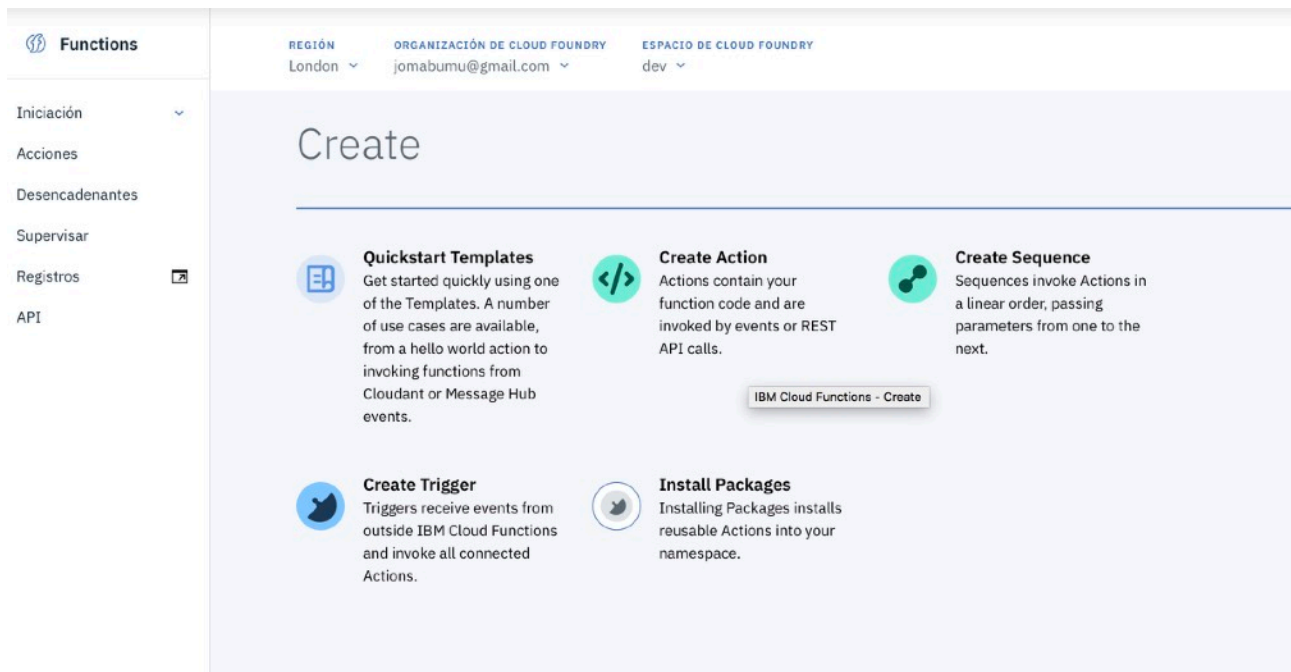
Subir

<input type="checkbox"/>	Nombre de objeto	Archivado	Tamaño	Última modificación
<input type="checkbox"/>	result/jobid=08a084c2-8188-4625-9756-4e9ef320ff29		0 bytes	18/05/2019 5:52:25 P. M. ***
<input type="checkbox"/>	result/jobid=08a084c2-8188-4625-9756-4e9ef320ff29/_SUCCESS		0 bytes	18/05/2019 5:52:28 P. M. ***
<input type="checkbox"/>	result/jobid=08a084c2-8188-4625-9756-4e9ef320ff29/part-00000-b46673a1-7668-4c83-81d1-4b4810b9bd74-c000-attempt_20190518155227_0029_m_000000_0.csv		270,7 KB	18/05/2019 5:52:28 P. M. ***
<input type="checkbox"/>	result/jobid=20f4eb54-0aff-4019-a0f4-4649478083e8		0 bytes	18/05/2019 6:04:53 P. M. ***
<input type="checkbox"/>	result/jobid=20f4eb54-0aff-4019-a0f4-4649478083e8/_SUCCESS		0 bytes	18/05/2019 6:04:57 P. M. ***
<input type="checkbox"/>	result/jobid=20f4eb54-0aff-4019-a0f4-4649478083e8/part-00000-05e82349-f4db-47a3-9114-0b80155695b8-c000-attempt_20190518160456_0008_m_000000_0.csv		8,4 KB	18/05/2019 6:04:57 P. M. ***
<input type="checkbox"/>	result/jobid=296ec545-4651-4ff6-9df1-f97bb763777f		0 bytes	18/05/2019 5:53:24 P. M. ***

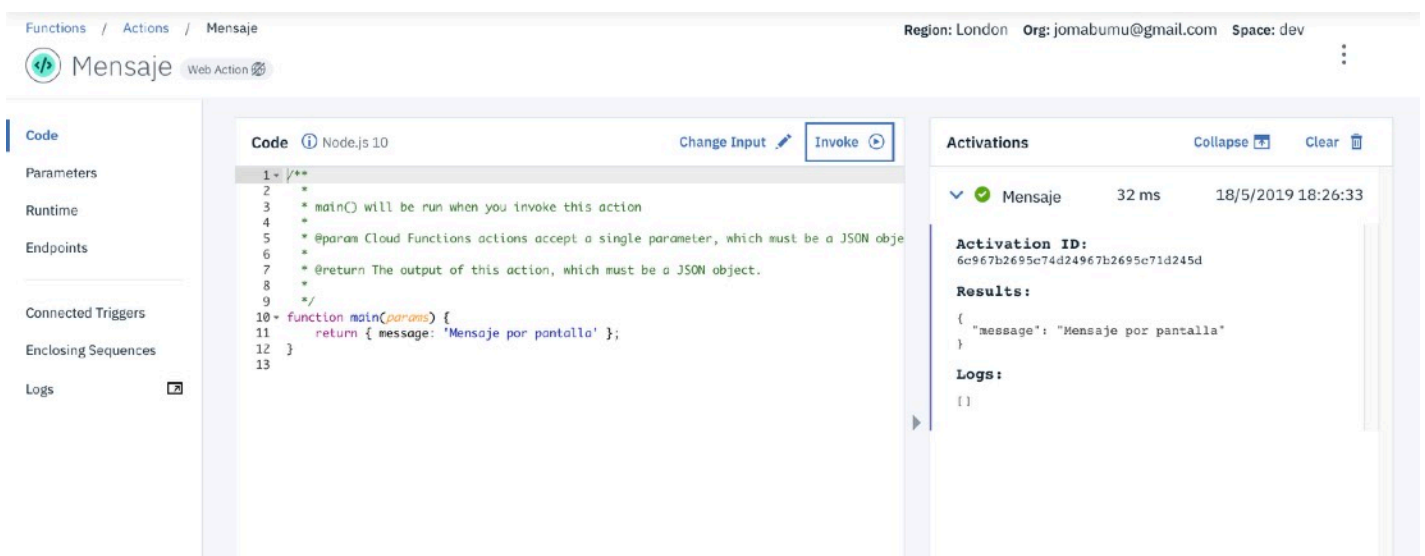
Realice las siguientes acciones documentando breve y concisamente el proceso:

5. Implemente una acción utilizando las funciones cloud que imprima un mensaje por pantalla cada vez que se ejecute. Compruebe la ejecución correcta de la acción.

Accedemos al servicio de las Cloud Functions dentro del catálogo de IBM, y ahí nos aparecen las opciones para poder crear acciones o triggers.



Generamos la acción, para mostrar un mensaje por pantalla:



6. Defina un trigger periódico que se ejecute cada 5 minutos. Relacione este trigger con la acción implementada en el apartado anterior para que la acción se ejecute cada 5 minutos.

Creamos un trigger, y le añadimos sentencia CRON para que se ejecute cada 5 minutos:

Functions

REGION: London ORGANIZACIÓN DE CLOUD FOUNDRY: jomabumu@gmail.com ESPACIO DE CLOUD FOUNDRY: dev

Connect Trigger

Periodic
Trigger an action based on time.
[Learn more about Periodic Triggers](#)

New Trigger Configuration

Trigger Name
Triggermensaje

Timer Settings
Specify UTC times as patterns or cron expression.

Cron Expression
Input a cron expression

Periodic
Trigger an action based on time.
[Learn more about Periodic Triggers](#)

Next time this trigger will fire:
sáb, 18/5/2019 18:35:00 CEST
sáb, 18/5/2019 18:40:00 CEST
sáb, 18/5/2019 18:45:00 CEST
sáb, 18/5/2019 18:50:00 CEST
sáb, 18/5/2019 18:55:00 CEST

New Trigger Configuration

Trigger Name
Triggermensaje

Timer Settings
Specify UTC times as patterns or cron expression.

Cron Expression
*/5 * * * *

JSON Payload ⓘ

Por último, conectamos la acción y el trigger generados:

Functions / Triggers / Triggermensaje

Region: London Org: jomabumu@gmail.com Space: dev

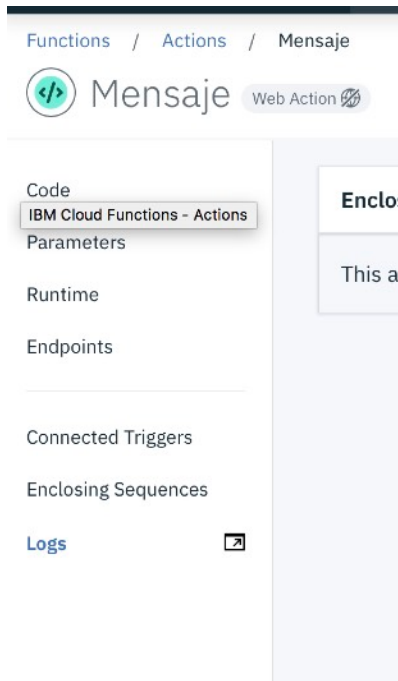
Triggermensaje

Connected Actions ⓘ [Add](#) +

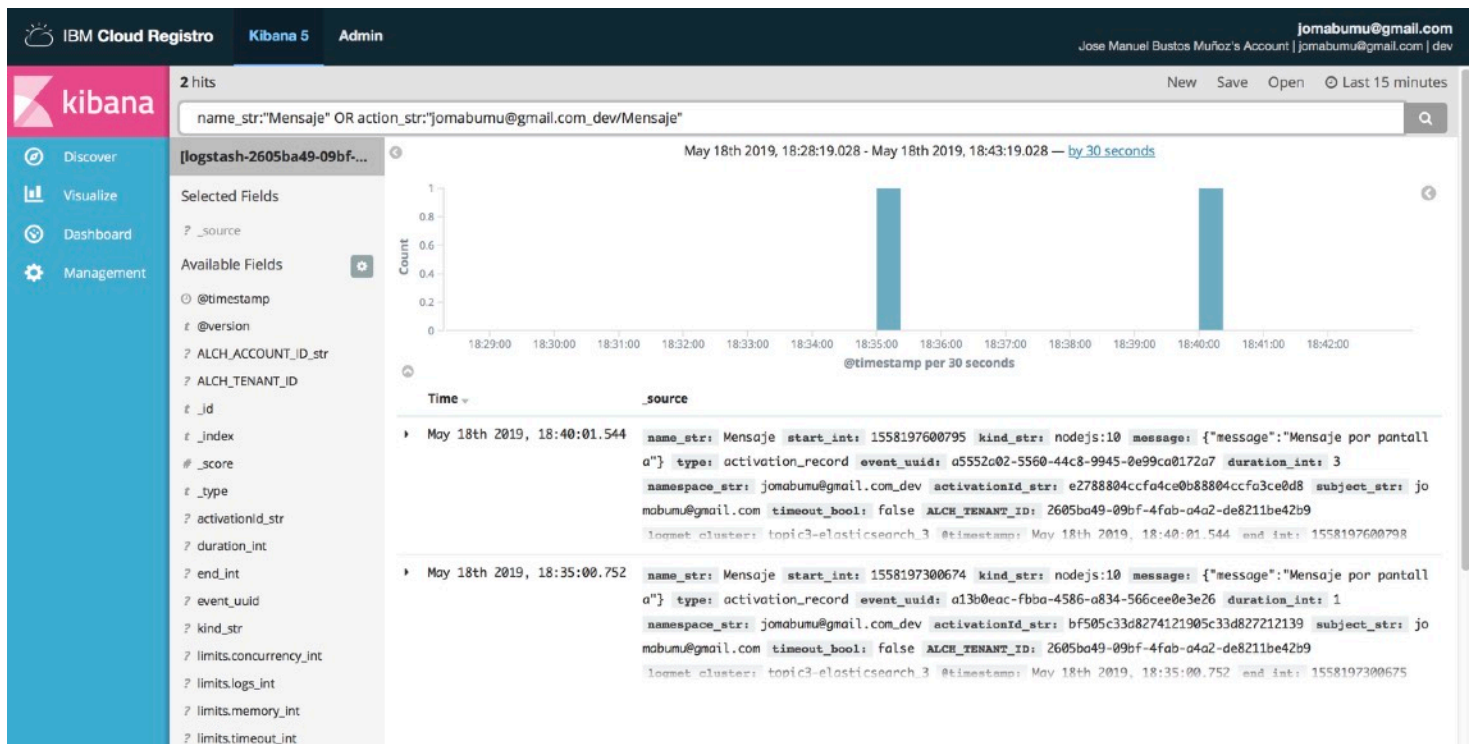
NAME	PACKAGE	CONNECTION
Mensaje	Default Package	<input checked="" type="checkbox"/> Enabled

7. Desarrolle un dashboard para monitorizar la ejecución de la acción definida. Justifique los elementos de visualización y / o análisis incluidos en el dashboard e incluya una captura del mismo en la memoria entregable.

En la parte izquierda de la pantalla hay un enlace abajo de “Logs” que nos abre una ventana del servicio de Kibana.



Al acceder al panel de KIBANA se pueden ver las dos primeras ejecuciones de la acción que enlazamos con el trigger, que la ha disparado a las 18:35 y a las 18:40, y se pueden ver los datos de estas acciones ejecutadas:



Aplique SWOT analysis para justificar las ventajas e inconvenientes potenciales derivadas del empleo de PaaS y procesos como los realizados en los ejercicios anteriores a proyectos como los que realice o haya realizado previamente:

8. Identifique y contextualice los aspectos principales a considerar sobre los proyectos.

El análisis SWOT identifica las fuerzas, debilidades, oportunidades y amenazas del elemento analizado. Las fuerzas y debilidades se suelen centrar en cuestiones internas y la experiencia adquirida, y el análisis de las oportunidades y amenazas se ocupa de aspectos externos y está orientado hacia el futuro.

Fortalezas:

- Comodidad. No hay que preocuparse de actualizaciones, ni preocuparse por la capacidad de almacenamiento o si el balanceo de carga es el adecuado.
- Menos inversión inicial y menos riesgo.
- Reducción de costes. Solo se paga por lo que se necesite, que se va estructurando de forma dinámica.
- Soporte ágil y rápido.
- Mayor disponibilidad y seguridad de los datos.

Debilidades:

- Nivel de confianza bajo en la seguridad de los datos. Si hay datos críticos para una empresa en Cloud genera desconfianza.
- Integración con el resto de aplicaciones del sistema de la empresa. Esto puede aumentar la complejidad para la empresa en cuestión.
- Sensación de cautividad del cliente.
- Posible incumplimiento de los acuerdos sobre el nivel de servicio. No sabemos del todo si se cumplirá el servicio con el nivel adecuado.
- Necesidad de disponer de los datos de la nube. Es importante que en todo momento tengamos acceso y dispongamos de los datos, o puede ser un gran inconveniente.

Oportunidades:

- Al ahorrar costes es una buena opción para paliar posibles crisis económicas.
- Es un mercado en auge y con expansión actualmente.
- El incremento del ancho de banda, la estabilidad, y la seguridad de Internet hace que este tipo de soluciones estén en auge y cada vez sean una mejor opción.

Amenazas:

- Un posible abaratamiento de los costes de infraestructura y mantenimiento más tradicionales.
- Aunque esté en auge, el mercado sigue siendo volátil y el auge de otra tendencia podría afectarle.
- Fallos importantes en cuestiones como seguridad, privacidad o disponibilidad que pueden hacer el auge de un competidor o tendencia contraria.