



PRÁCTICA ELASTICSEARCH

Sistemas Distribuidos

Práctica Elasticsearch

Desplegar una base de datos Elasticsearch en la nube y almacenar datos de una API pública para su posterior análisis.

Jaime Arana, Javier Álvarez

Práctica Elasticsearch 1.1 – Desplegar una base de datos en la nube

En este apartado creamos una cuenta de prueba por 14 días, del 01/04/2022 al 14/04/2022. Con esta cuenta vamos a poder desplegar de una base de datos Elasticsearch en la nube.

Práctica Elasticsearch 1.2 – Crear un usuario de servicio para nuestra base de datos

Una vez desplegada nuestra base de datos, creamos un usuario con role de superusuario y así poder hacer la ingesta de datos más adelante. Estas dos secciones de la parte están bien documentadas y son fáciles de seguir, por lo tanto, no se adjuntan capturas del progreso.

Práctica Elasticsearch 1.3 – Revisa el formato del API para crear el mapping de los datos

En este apartado se procede a configurar la base de datos, atendiendo a los datos que se van a ingestar más adelante. Dichos datos están publicados en un archivo XML y contienen el tráfico en tiempo real de la comunidad de Madrid.

Crear índice “tráfico-madrid”



The screenshot shows a REST client interface with the following details:


- URL:** `https://practica-elastic-search.es.us-central1.gcp.cloud.es.io:9243/tráfico-madrid`
- Method:** PUT
- Body (JSON):**

```
{
  "settings": {
    "number_of_shards": 3,
    "number_of_replicas": 1
  }
}
```
- Response (JSON):**

```
{
  "acknowledged": true,
  "shards_acknowledged": true,
  "index": "tráfico-madrid"
}
```
- Status:** 200 OK, Time: 840 ms, Size: 399 B

Actualizar los settings

Body Cookies Headers (8) Test Results  Status: 200 OK Time: 840 ms Size: 399 B [Save Response](#) 



Pretty Raw Preview Visualize JSON 


```
1 {
2   "acknowledged": true,
3   "shards_acknowledged": true,
4   "index": "trafico-madrid"
5 }
```

Comprobar la modificación de los settings

GET https://practica-elastic-search.es.us-central1.gcp.cloud.es.io:9243/trafico-madrid/_settings [Send](#)

Params Authorization Headers (7) Body Pre-request Script Tests Settings Cookies

Body Cookies Headers (8) Test Results  Status: 200 OK Time: 137 ms Size: 575 B [Save Response](#) 








Pretty Raw Preview Visualize JSON 

```
1 {
2   "trafico-madrid": {
3     "settings": {
4       "index": {
5         "routing": {
6           "allocation": {
7             "include": {
8               "_tier_preference": "data_content"
9             }
10          }
11        },
12        "mapping": {
13          "total_fields": {
14            "limit": "40"
15          }
16        },
17        "number_of_shards": "3",
18        "provided_name": "trafico-madrid",
19        "creation_date": "1648805063025",
20        "number_of_replicas": "1",
21        "uuid": "X4wbPyU8Q00FtnVL7JAqww",
22        "version": 8,
23        "created": "8010299"
24      }
25    }
26  }
27 }
28 }
```

Crear mapping para almacenar el modelo de datos

PUT https://practica-elastic-search.es.us-central1.gcp.cloud.es.io:9243/trafico-madrid/_mapping [Send](#)

Params Authorization Headers (9) Body Pre-request Script Tests Settings Cookies

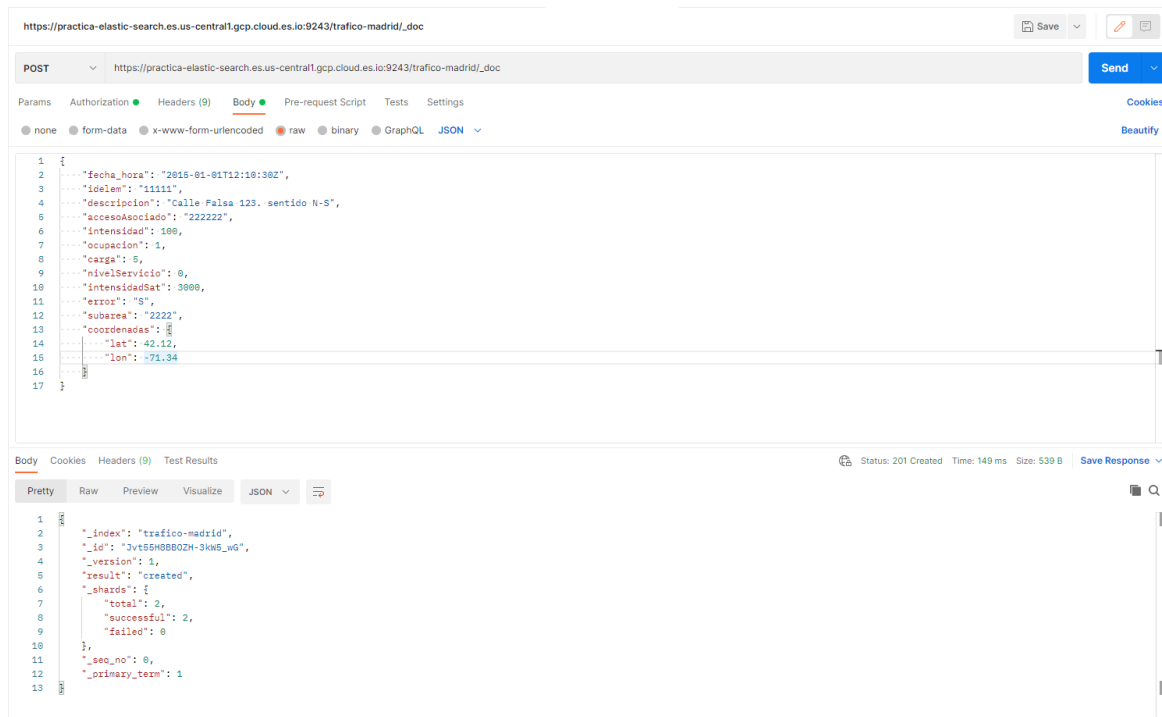
 none  form-data  x-www-form-urlencoded  raw  binary  GraphQL JSON 

[Beautify](#)

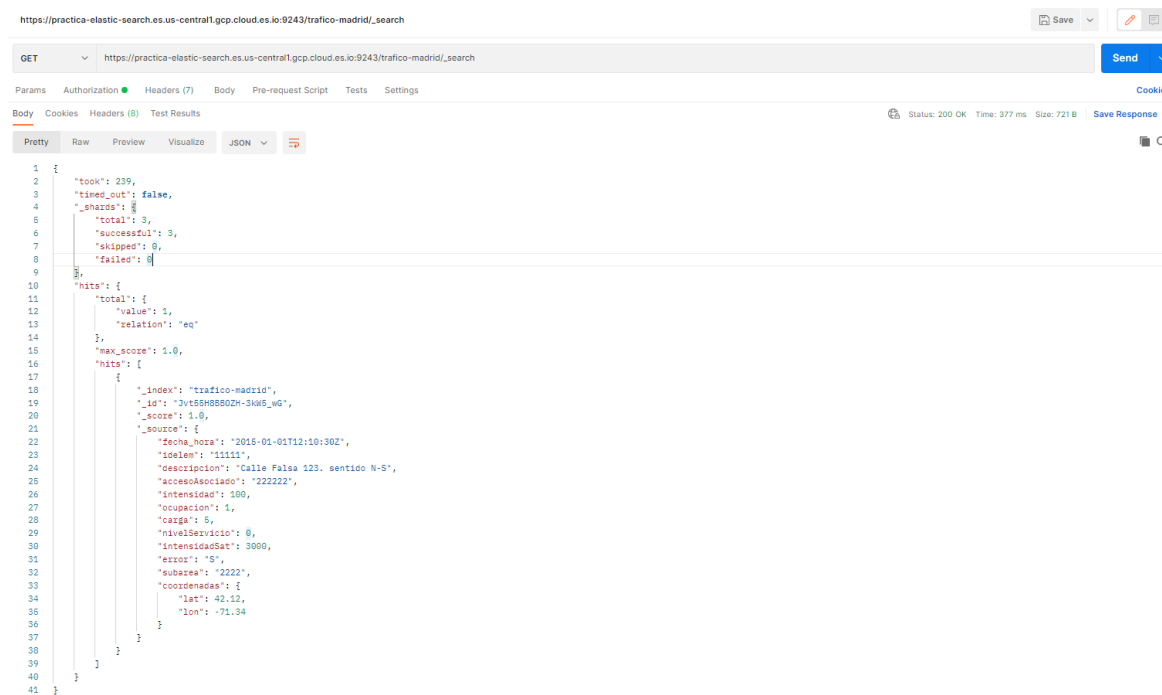
```
1 {
2   "properties": {
3     "fecha_hora": {
4       "type": "date"
5     },
6     "idoleon": {
7       "type": "text"
8     },
9     "descripcion": {
10      "type": "text"
11    },
12    "accesoAsociado": {
13      "type": "text"
14    },
15    "intensidad": {
16      "type": "long"
17    },
18    "ocupacion": {
19      "type": "long"
20    },
21    "carga": {
22      "type": "long"
23    },
24    "nivelServicio": {
25      "type": "long"
26    },
27    "intensidadSat": {
28      "type": "long"
29    },
30    "error": {
31      "type": "text"
32    },
33    "subarea": {
34      "type": "text"
35    },
36    "coordenadas": {
37      "type": "geo_point"
38    }
39  }
40 }
```



Añadir una prueba de datos manualmente



Recuperar la entrada de prueba de datos



Borrar la entrada de prueba

SSDD_Elastic_Search / https://practica-elastic-search.es.us-central1.gcp.cloud.es.io:9243/trafico-madrid/_search Save ... 🔗 📄

DELETE ▼ https://practica-elastic-search.es.us-central1.gcp.cloud.es.io:9243/trafico-madrid/_doc/Jvt55H8B80ZH-3kW5_wG Send ▼

Params Authorization ● Headers (7) Body Pre-request Script Tests Settings Cookies

● none ● form-data ● x-www-form-urlencoded ● raw ● binary ● GraphQL

This request does not have a body

Body Cookies Headers (8) Test Results 🌐 Status: 200 OK Time: 216 ms Size: 480 B Save Response ▼

Pretty Raw Preview Visualize JSON ▼ 🔍

```
1  {
2    "_index": "trafico-madrid",
3    "_id": "Jvt55H8B80ZH-3kW5_wG",
4    "_version": 2,
5    "result": "deleted",
6    "_shards": {
7      "total": 2,
8      "successful": 2,
9      "failed": 0
10   },
11    "_seq_no": 1,
12    "_primary_term": 1
13  }
```

Práctica Elasticsearch 1.4 y 1.6

Los siguientes apartados de la práctica se completan el notebook adjuntado en la entrega.

Práctica Elasticsearch 1.5

Si entramos en la interfaz gráfica web para ver las entradas de datos, podemos observar que se han añadido campos que no hemos definido en el mapping. Estos campos pueden ser velocidad, por ejemplo. Esto no es ningún problema para la base de datos de Elasticsearch, ya que permite el mapping dinámico. Esto significa que Elasticsearch puede añadir campos nuevos automáticamente.