PRÁCTICA ELASTICSEARCH HADOOP

PARTE 1 - Configuración ES-Hadoop

SSH en el navegador

```
Welcome to Ubuntu 22.04.5 LTS (GNU/Linux 6.8.0-1025-gcp x86_64)

 * Documentation:  https://help.ubuntu.com
 * Management:     https://landscape.canonical.com
 * Support:        https://ubuntu.com/pro

 System information as of Fri Mar 14 19:17:00 UTC 2025

  System load:  2.03               Processes:              180
  Usage of /:   30.9% of 48.27GB   Users logged in:        0
  Memory usage: 35%                IPv4 address for ens4: 10.154.0.6
  Swap usage:   0%

 * Strictly confined Kubernetes makes edge and IoT secure. Learn how MicroK8s
   just raised the bar for easy, resilient and secure K8s cluster deployment.

   https://ubuntu.com/engage/secure-kubernetes-at-the-edge

Expanded Security Maintenance for Applications is not enabled.

0 updates can be applied immediately.

Enable ESM Apps to receive additional future security updates.
See https://ubuntu.com/esm or run: sudo pro status

New release '24.04.2 LTS' available.
Run 'do-release-upgrade' to upgrade to it.


Last login: Fri Mar 14 18:35:05 2025 from 35.235.248.241
leticm03@practica-hadoop-elastic-m:~$ gsutil cp gs://bucket-practica-elastic/jars/elastic/elasticsearch-hadoop-8.14.1.jar .
CommandException: No URLs matched: gs://bucket-practica-elastic/jars/elastic/elasticsearch-hadoop-8.14.1.jar
leticm03@practica-hadoop-elastic-m:~$ gsutil cp gs://bucket-practica-elastic/elasticsearch-hadoop-8.14.1.jar .
Copying gs://bucket-practica-elastic/elasticsearch-hadoop-8.14.1.jar...
/ [1 files][  2.1 MiB/  2.1 MiB]
Operation completed over 1 objects/2.1 MiB.
leticm03@practica-hadoop-elastic-m:~$ gsutil cp gs://bucket-practica-elastic/commons-httpclient-3.1.jar .
Copying gs://bucket-practica-elastic/commons-httpclient-3.1.jar...
/ [1 files][297.8 KiB/297.8 KiB]
Operation completed over 1 objects/297.8 KiB.
```

PARTE 2 – Configuración server ElasticSearch

```
#----------------------- BEGIN SECURITY AUTO CONFIGURATION -----------------------
#
# The following settings, TLS certificates, and keys have been automatically
# generated to configure Elasticsearch security features on 14-03-2025 20:19:31
#
# --------------------------------------------------------------------------------

# Enable security features
xpack.security.enabled: false

xpack.security.enrollment.enabled: true

# Enable encryption for HTTP API client connections, such as Kibana, Logstash, and Agents
xpack.security.http.ssl:
  enabled: false
  keystore.path: certs/http.p12

# Enable encryption and mutual authentication between cluster nodes
xpack.security.transport.ssl:
  enabled: true
  verification_mode: certificate
  keystore.path: certs/transport.p12
  truststore.path: certs/transport.p12
# Create a new cluster with the current node only
# Additional nodes can still join the cluster later
cluster.initial_master_nodes: ["instance-elastic"]

# Allow HTTP API connections from anywhere
# Connections are encrypted and require user authentication
http.host: 0.0.0.0

# Allow other nodes to join the cluster from anywhere
# Connections are encrypted and mutually authenticated
#transport.host: 0.0.0.0

#----------------------- END SECURITY AUTO CONFIGURATION -----------------------
```

```
^G Get Help    ^O Write Out    ^W Where Is    ^K Cut Text    ^J Justify    ^C Cur
^X Exit        ^R Read File    ^\ Replace     ^U Paste Text  ^T To Spell   ^  Go
```

PARTE 3 – Configuración en Cluster Hadoop de Conexión con ES

```
Enable ESM Apps to receive additional future security updates.
See https://ubuntu.com/esm or run: sudo pro status

New release '24.04.2 LTS' available.
Run 'do-release-upgrade' to upgrade to it.


Last login: Fri Mar 14 20:41:42 2025 from 35.235.248.242
leticm03@practica-hadoop-elastic-m:~$ sudo sed -i '$d' /etc/hive/conf.dist/hive-site.xml
leticm03@practica-hadoop-elastic-m:~$ sudo sed -i '$a \  <property>\n    <name>es.nodes</name>\n    <value
>34.105.248.162</value>\n  </property>\n' /etc/hive/conf.dist/hive-site.xml
leticm03@practica-hadoop-elastic-m:~$ sudo sed -i '$a \  <property>\n    <name>es.port</name>\n    <value>
9200</value>\n  </property>\n' /etc/hive/conf.dist/hive-site.xml
leticm03@practica-hadoop-elastic-m:~$ sudo sed -i '$a \  <property>\n    <name>es.nodes.wan.only</name>\n
   <value>true</value>\n  </property>\n' /etc/hive/conf.dist/hive-site.xml
leticm03@practica-hadoop-elastic-m:~$ sudo sed -i '$a \  <property>\n    <name>hive.aux.jars.path</name>\n
    <value>/usr/lib/hive/lib/elasticsearch-hadoop-8.14.1.jar,/usr/lib/hive/lib/commons-httpclient-3.1.jar</
value>\n  </property>\n</configuration>' /etc/hive/conf.dist/hive-site.xml
leticm03@practica-hadoop-elastic-m:~$ sudo cp elasticsearch-hadoop-8.14.1.jar /usr/lib/hive/lib/
leticm03@practica-hadoop-elastic-m:~$ sudo cp commons-httpclient-3.1.jar /usr/lib/hive/lib/
leticm03@practica-hadoop-elastic-m:~$ sudo service hive restart
Failed to restart hive.service: Unit hive.service not found.
leticm03@practica-hadoop-elastic-m:~$ sudo systemctl restart hive-server2
leticm03@practica-hadoop-elastic-m:~$
```

```
  <property>
    <name>es.port</name>
    <value>9200</value>
  </property>

  <property>
    <name>es.nodes.wan.only</name>
    <value>true</value>
  </property>

  <property>
    <name>hive.aux.jars.path</name>
    <value>/usr/lib/hive/lib/elasticsearch-hadoop-8.14.1.jar,/usr/lib/hive/lib/commons-httpclient-3.1.jar</value>
  </property>
</configuration>
leticm03@practica-hadoop-elastic-m:~$
```

PARTE 4 – Conectar datos

```
Last login: Mon Mar 17 11:06:02 2025 from 35.235.242.2
leticm03@practica-hadoop-elastic-m:~$ curl -X POST "http://35.242.138.80:9200/_bulk" -H 'Content-Type: applicat
ion/json' -d'
{ "index": { "_index": "alumnos", "_id": "3" } }
{ "id": 3, "name": "Carlos", "last_name": "González" }
{ "index": { "_index": "alumnos", "_id": "4" } }
{ "id": 4, "name": "María", "last_name": "López" }
{ "index": { "_index": "alumnos", "_id": "5" } }
{ "id": 5, "name": "Luis", "last_name": "Martínez" }
{ "index": { "_index": "alumnos", "_id": "7" } }
{ "id": 7, "name": "Sofía", "last_name": "Ramírez" }
{ "index": { "_index": "alumnos", "_id": "8" } }
{ "id": 8, "name": "Pedro", "last_name": "Hernández" }
'
{"errors":false,"took":82,"items":[{"index":{"_index":"alumnos","_id":"3","_version":1,"result":"created","_sha
rds":{"tsion":1,"result":"created","_shards":{"total":2,"successful":1,"failed":0},"_seq_no":2,"_primary_term":
1,"status":201}}q_no":3,"_primary_term":1,"status":201}},{"index":{"_index":"alumnos","_id":"7","_version":1,"r
esult":"created","_shard
```

```
leticm03@practica-hadoop-elastic-m:~$ curl -X GET "http://35.242.138.80:9200/alumnos/_search?pretty"
{
  "took" : 13,
  "timed_out" : false,
  "_shards" : {
    "total" : 1,
    "successful" : 1,
    "skipped" : 0,
    "failed" : 0
  },
  "hits" : {
    "total" : {
      "value" : 6,
      "relation" : "eq"
    },
    "max_score" : 1.0,
    "hits" : [
      {
        "_index" : "alumnos",
        "_id" : "6",
        "_score" : 1.0,
        "_source" : {
          "title" : "New Document",
          "content" : "This is a new document for the master class",
          "tag" : [
            "general",
            "testing"
          ]
        }
      },
      {
        "_index" : "alumnos",
        "_id" : "3",
        "_score" : 1.0,
        "_source" : {
          "id" : 3,
          "name" : "Carlos",
          "last_name" : "González"
        }
      },
      {
        "_index" : "alumnos",
        "_id" : "4",
        "_score" : 1.0,
        "_source" : {
          "id" : 4,
          "name" : "María",
          "last_name" : "López"
        }
      },
      {
        "_index" : "alumnos",
        "_id" : "5",
        "_score" : 1.0,
        "_source" : {
          "id" : 5,
          "name" : "Luis",
          "last_name" : "Martínez"
```

```
          "last_name" : "Martínez"
        }
      },
      {
        "_index" : "alumnos",
        "_id" : "7",
        "_score" : 1.0,
        "_source" : {
          "id" : 7,
          "name" : "Sofía",
          "last_name" : "Ramírez"
        }
      },
      {
        "_index" : "alumnos",
        "_id" : "8",
        "_score" : 1.0,
        "_source" : {
          "id" : 8,
          "name" : "Pedro",
          "last_name" : "Hernández"
        }
      }
    ]
  }
}
leticm03@practica-hadoop-elastic-m:~$
```

PARTE 5 - Kibana