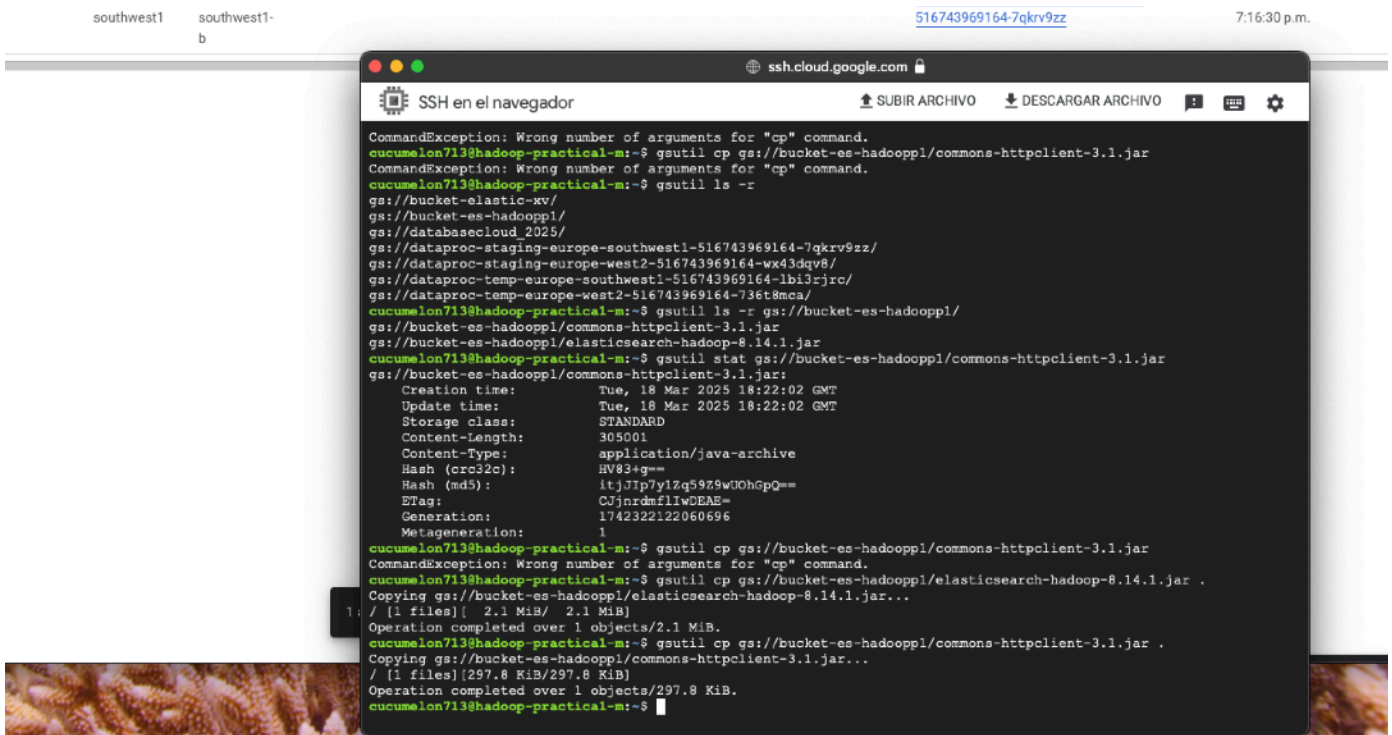


## PART 1 - Configuration ElasticSearch-Hadoop (ES-Hadoop)

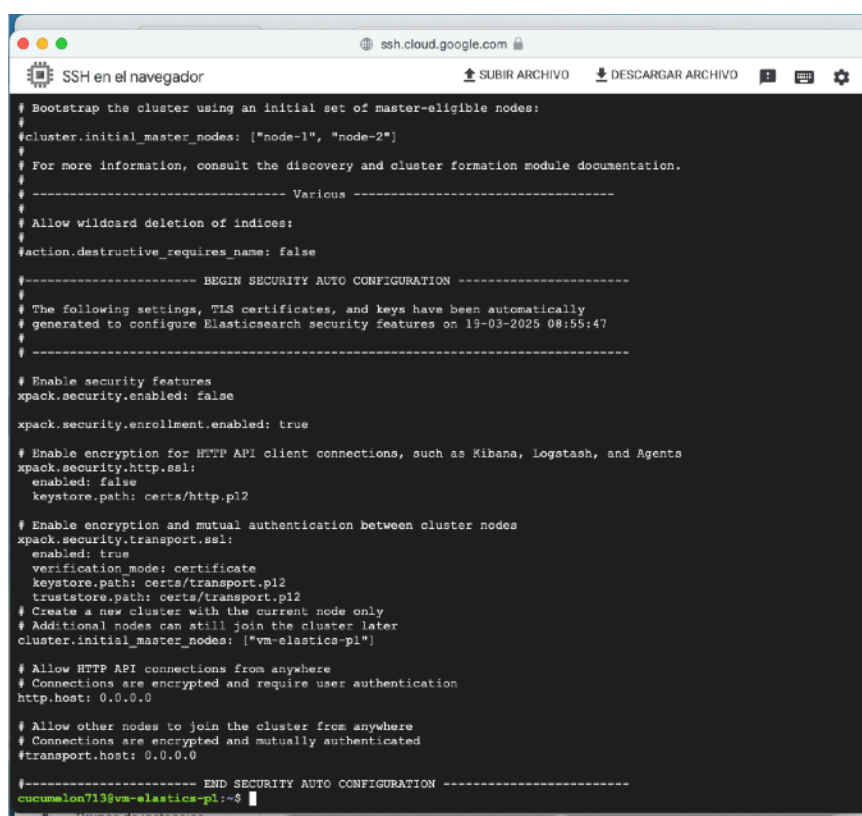
Screenshot of SSH terminal from Hadoop cluster after Hadoop has been configured with elasticsearch-hadoop and commons-httpclient.



```
CommandException: Wrong number of arguments for "cp" command.
cucumelon713@hadoop-practical-m:~$ gsutil cp gs://bucket-es-hadooppl/commons-httpclient-3.1.jar
CommandException: Wrong number of arguments for "cp" command.
cucumelon713@hadoop-practical-m:~$ gsutil ls -r
gs://bucket-elastic-xv/
gs://bucket-es-hadooppl/
gs://databasecloud_2025/
gs://dataproc-staging-europe-southwest1-516743969164-7qkrv9zz/
gs://dataproc-staging-europe-west2-516743969164-wx43dqv8/
gs://dataproc-temp-europe-southwest1-516743969164-lbi3z-jrc/
gs://dataproc-temp-europe-west2-516743969164-736t8mea/
cucumelon713@hadoop-practical-m:~$ gsutil ls -r gs://bucket-es-hadooppl/
gs://bucket-es-hadooppl/commons-httpclient-3.1.jar
gs://bucket-es-hadooppl/elasticsearch-hadoop-8.14.1.jar
cucumelon713@hadoop-practical-m:~$ gsutil stat gs://bucket-es-hadooppl/commons-httpclient-3.1.jar
gs://bucket-es-hadooppl/commons-httpclient-3.1.jar:
  Creation time:      Tue, 18 Mar 2025 18:22:02 GMT
  Update time:       Tue, 18 Mar 2025 18:22:02 GMT
  Storage class:     STANDARD
  Content-Length:    305001
  Content-Type:      application/java-archive
  Hash (crc32c):     HV83+g==
  Hash (md5):        ItjJip7y1Zq59Z9wUOhSpQ==
  ETag:              CJjnrdfllwDEAE=
  Generation:        1742322122060696
  Metageneration:    1
cucumelon713@hadoop-practical-m:~$ gsutil cp gs://bucket-es-hadooppl/commons-httpclient-3.1.jar
CommandException: Wrong number of arguments for "cp" command.
cucumelon713@hadoop-practical-m:~$ gsutil cp gs://bucket-es-hadooppl/elasticsearch-hadoop-8.14.1.jar .
Copying gs://bucket-es-hadooppl/elasticsearch-hadoop-8.14.1.jar...
 / [1 files] [ 2.1 MiB / 2.1 MiB]
Operation completed over 1 objects/2.1 MiB.
cucumelon713@hadoop-practical-m:~$ gsutil cp gs://bucket-es-hadooppl/commons-httpclient-3.1.jar .
Copying gs://bucket-es-hadooppl/commons-httpclient-3.1.jar...
 / [1 files] [297.8 KiB / 297.8 KiB]
Operation completed over 1 objects/297.8 KiB.
cucumelon713@hadoop-practical-m:~$
```

## PART 2 - Configuration of ElasticSearch server

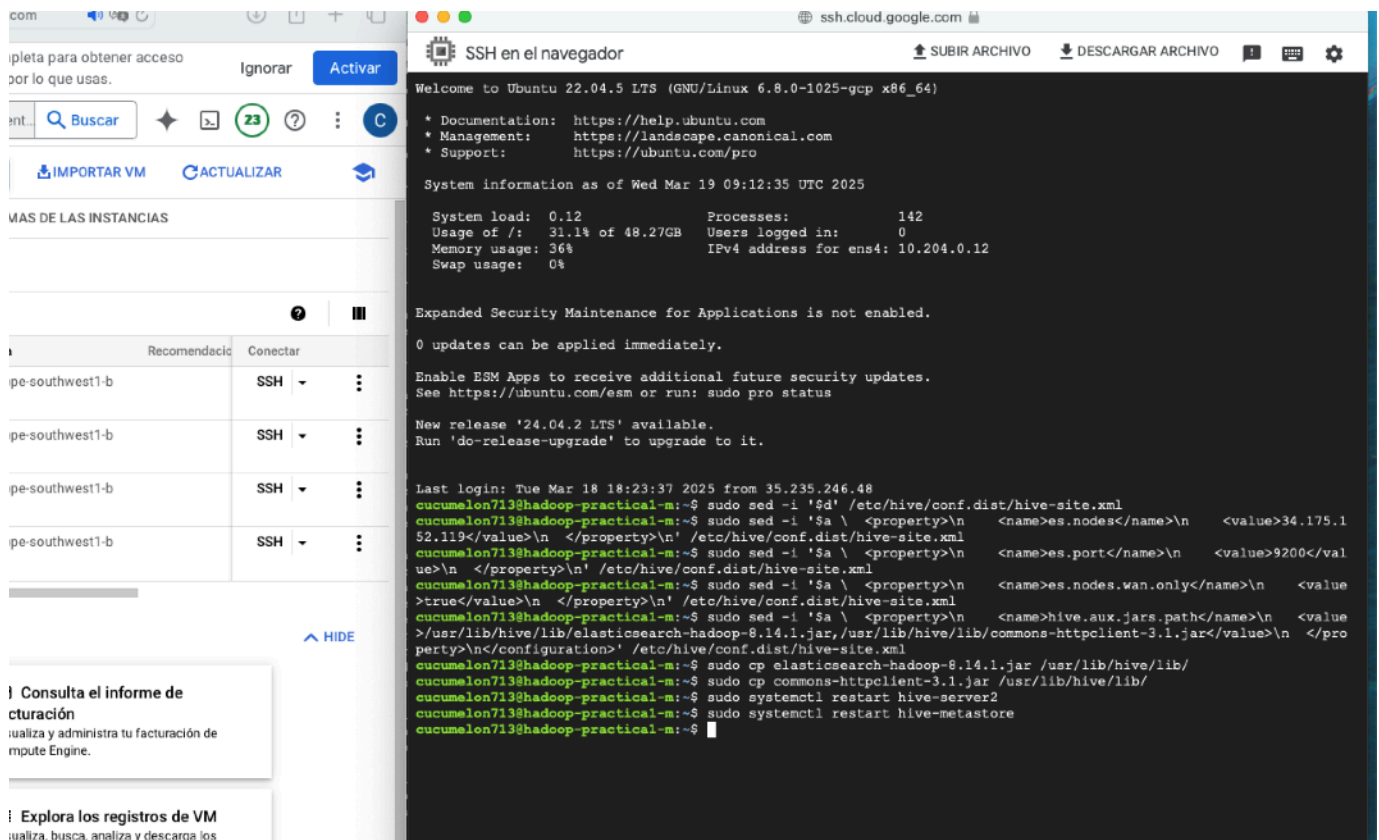
Screenshot of SSH terminal from ElasticSearch Instance where ElasticSearch has been configured.



```
# Bootstrap the cluster using an initial set of master-eligible nodes:
# cluster.initial_master_nodes: ["node-1", "node-2"]
# For more information, consult the discovery and cluster formation module documentation.
# ----- Various -----
# Allow wildcard deletion of indices:
# action.destructive_requires_name: false
# ----- BEGIN SECURITY AUTO CONFIGURATION -----
# The following settings, TLS certificates, and keys have been automatically
# generated to configure Elasticsearch security features on 19-03-2025 08:55:47
# -----
# 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: ["vm-elastics-pl"]
# 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 -----
cucumelon713@vm-elastics-pl:~$
```

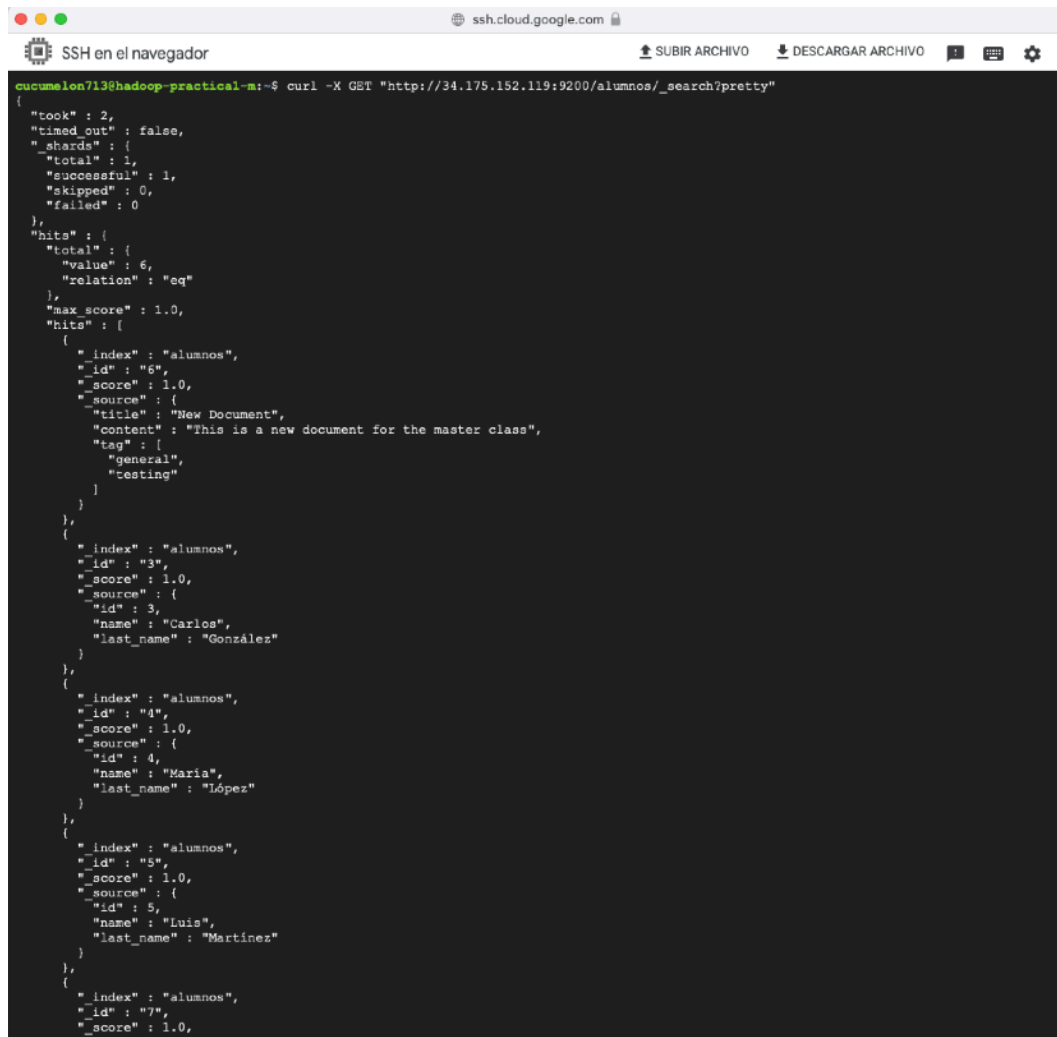
## PART 3 - Configuration in cluster Hadoop of connection with ES.

Screenshot of SSH terminal from Hadoop cluster to connect ES server with Hive.



## PART 4 - Connecting Data

Screenshot of the Hadoop cluster console with the query result.



The screenshot shows an SSH terminal window titled "SSH en el navegador" with a browser address bar displaying "ssh.cloud.google.com". The terminal prompt is "cucumalen713@hadoop-practical-m:~\$". The user has executed the command "curl -X GET 'http://34.175.152.119:9200/alumnos/\_search?pretty'", and the output is a JSON response from an Elasticsearch search API. The JSON indicates 6 hits, with the first three shown in detail. Each hit includes an index name, an ID, a score of 1.0, and a source object containing document details like title, content, and tags.

```
{
  "took": 2,
  "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": "Maria",
          "last_name": "López"
        }
      },
      {
        "_index": "alumnos",
        "_id": "5",
        "_score": 1.0,
        "_source": {
          "id": 5,
          "name": "Luis",
          "last_name": "Martínez"
        }
      },
      {
        "_index": "alumnos",
        "_id": "7",
        "_score": 1.0,

```



This block shows the continuation of the JSON output from the previous screenshot. It contains the last two hits from the search results and the closing brackets of the JSON object. The hits for IDs 7 and 8 are shown, with source objects containing names and last names.

```
    {
      "_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"
      }
    }
  ]
}
```

cucumalen713@hadoop-practical-m:~\$

PART 5 - KIBANA

Screenshot of the Kibana console with a simple visualisation.



Caption