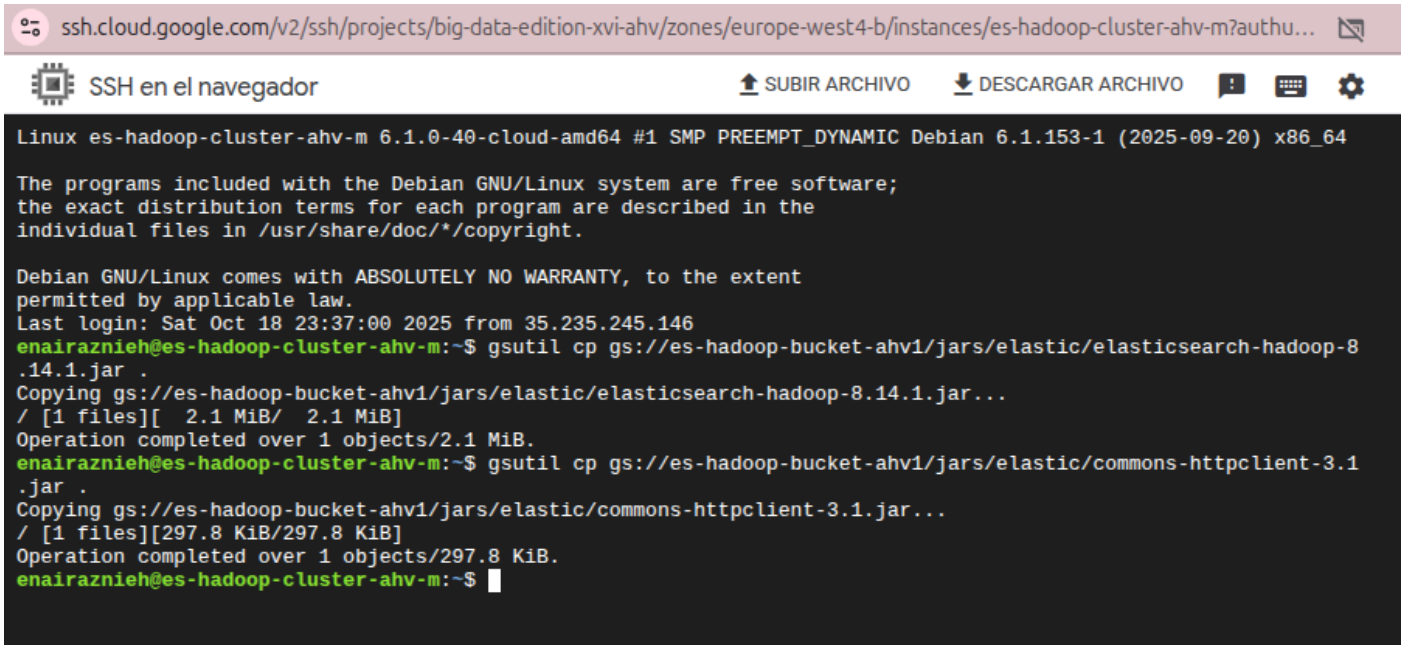
Big Data Architecture
Final practical deliverable

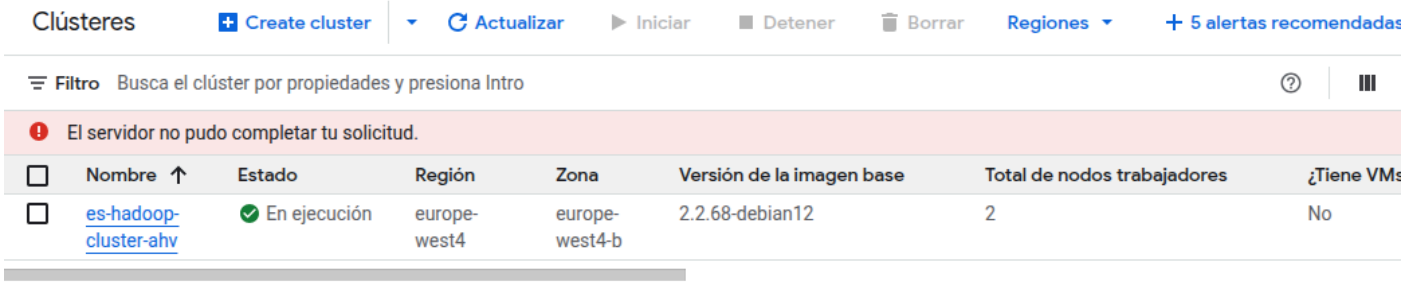
by Ariane (Ariadna) Heinz Vallribera

PART 1 SCREENSHOTS:

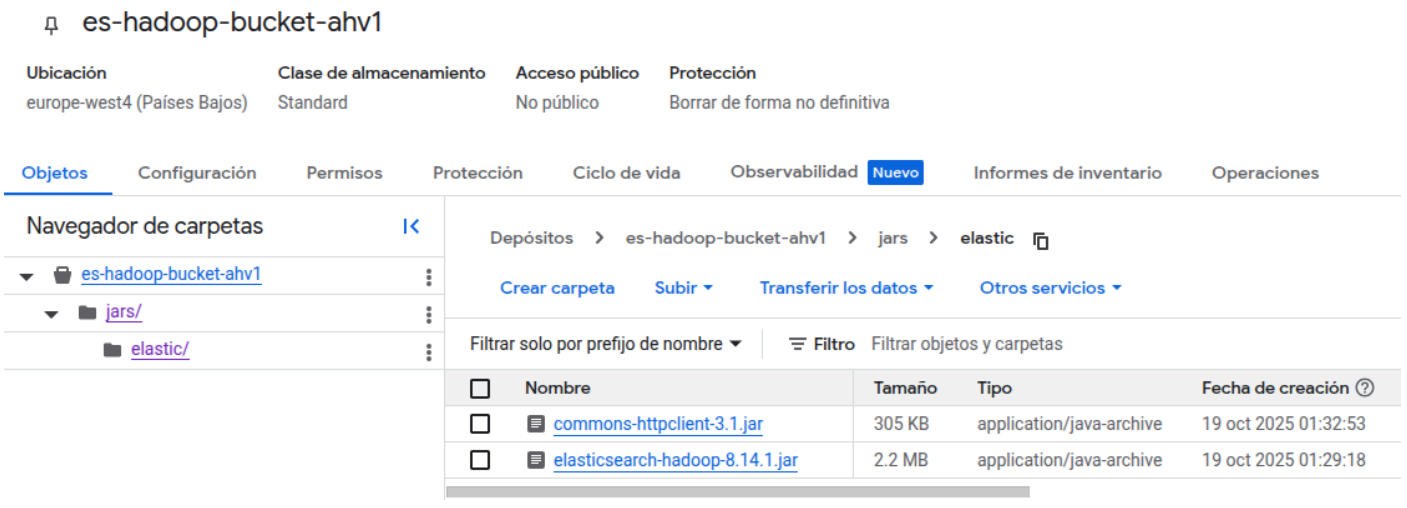
SSH console for the Hadoop cluster once the loading of the configuration files is complete:



Cluster with 2 worker nodes:

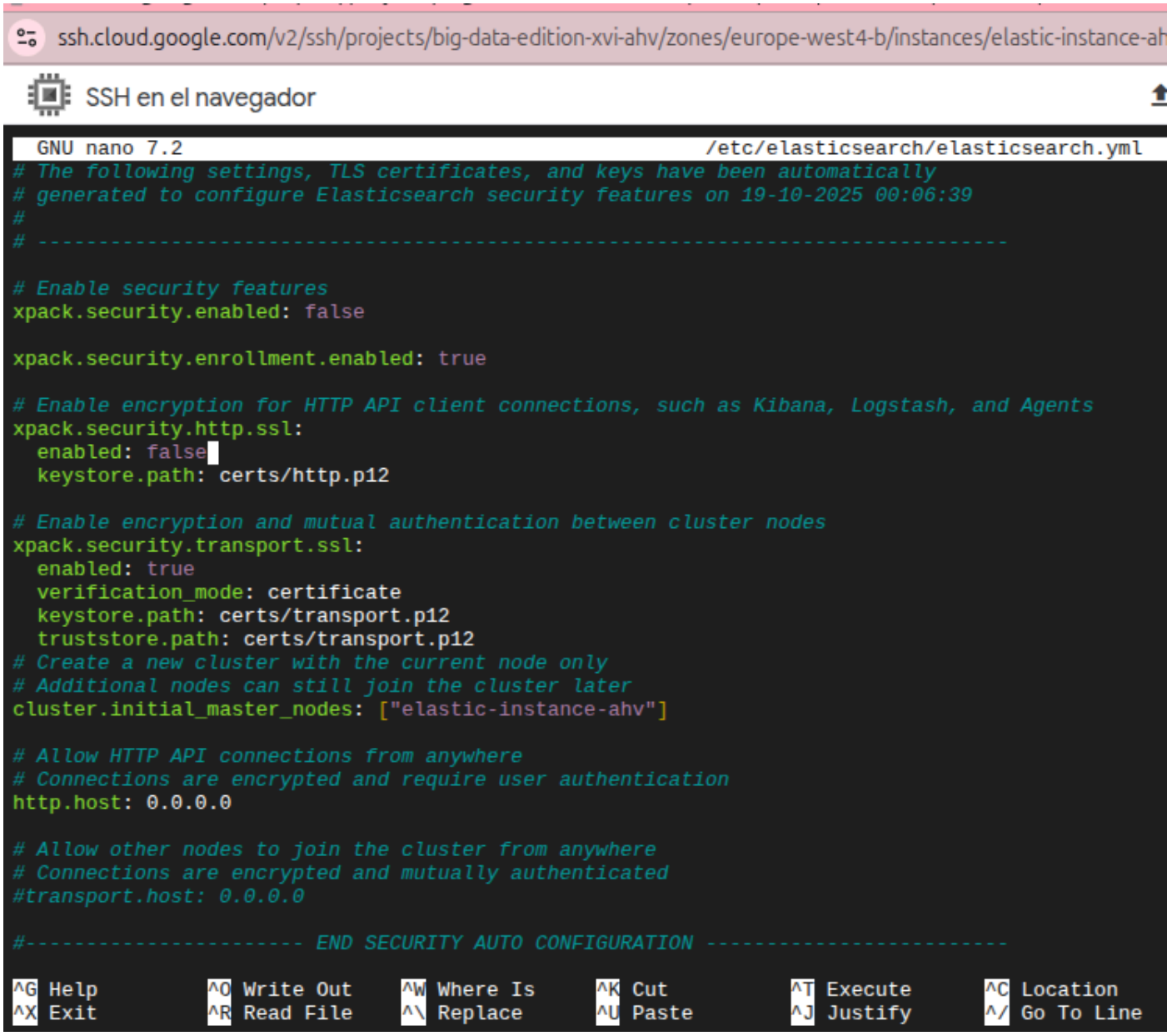


Bucket that contains configuration files:



PART 2 SCREENSHOTS:

Configuration changes in the elasticsearch.yml file:



Verification that the cluster can indeed access the Elastic machine:

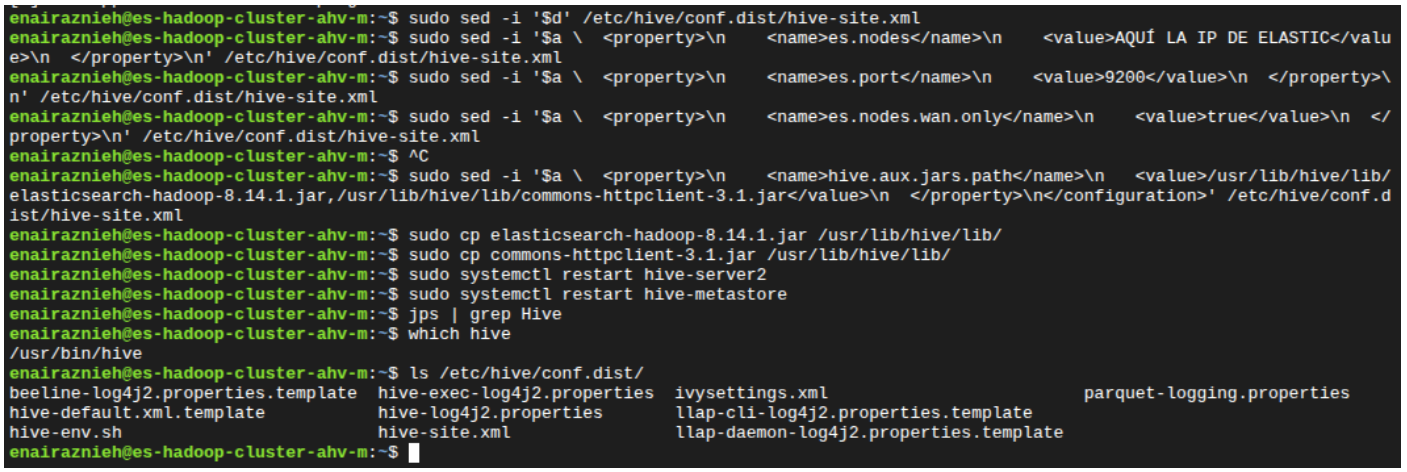
```
enairaznieh@es-hadoop-cluster-ahv-m:~$ ping 34.7.74.62
PING 34.7.74.62 (34.7.74.62) 56(84) bytes of data.
64 bytes from 34.7.74.62: icmp_seq=1 ttl=61 time=1.12 ms
64 bytes from 34.7.74.62: icmp_seq=2 ttl=61 time=0.738 ms
64 bytes from 34.7.74.62: icmp_seq=3 ttl=61 time=0.439 ms
64 bytes from 34.7.74.62: icmp_seq=4 ttl=61 time=0.645 ms
64 bytes from 34.7.74.62: icmp_seq=5 ttl=61 time=0.539 ms
```

Additional firewalls I created manually:

☐	elastic-kibana-firewall-1	Entrada	Aplicar a	Rangos de IP:	tcp:5601, 9200	Permitir	1000
☐	hadoop-firewall-1	Entrada	Aplicar a	Rangos de IP:	tcp:8088, 9870	Permitir	1000

PART 3 SCREENSHOTS:

Configuration process of connection with ES in Hadoop Cluster:



The screenshot also shows the three commands I ran to ensure that Hive is indeed in the es-hadoop cluster:

- The command `jps | grep Hive` resulted in no output, signifying that no hive-related java processes were running at the moment.
- The command `which hive` returned the location where the *hive* executable is located.
- The command `ls /etc/hive/conf.dist/` returned the list of files in the configuration directory of *hive*, indicating that they are indeed there.

PART 4 SCREENSHOTS:

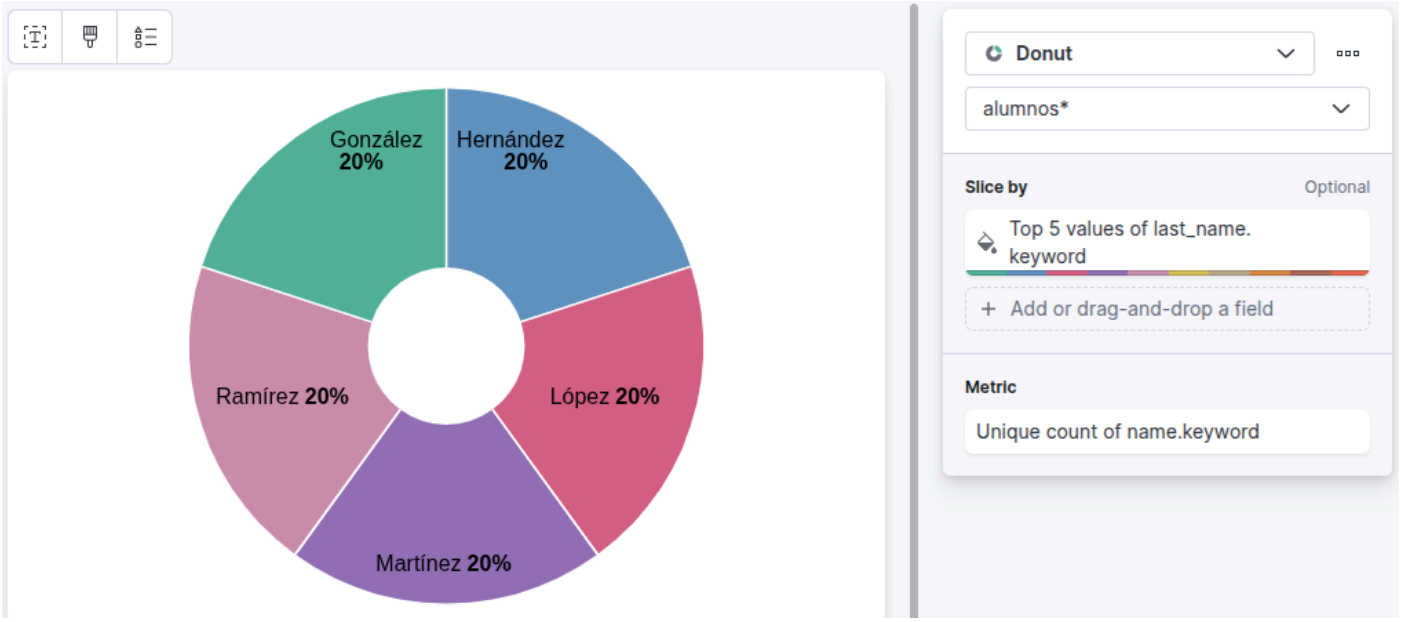
Result of query from Hadoop cluster:

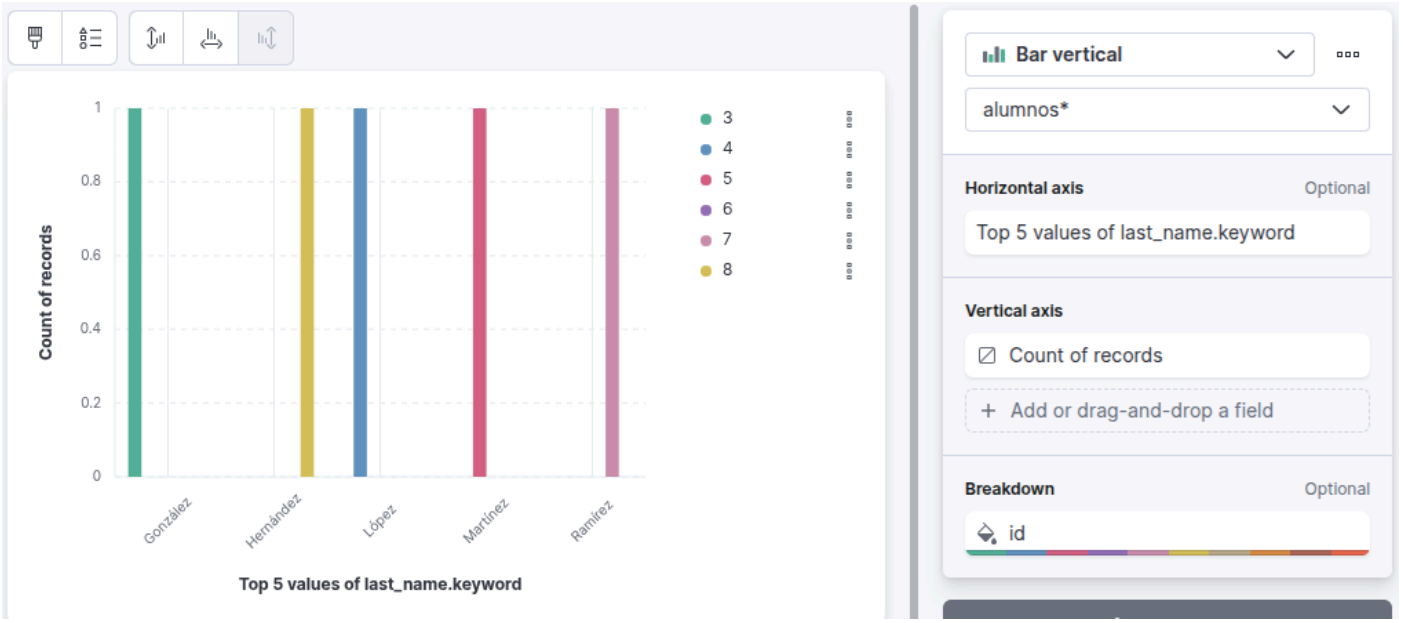
```
enairaznieh@es-hadoop-cluster-ahv-m:~$ curl -X GET "http://34.7.74.62:9200/alumnos/_search?pretty"
{
  "took" : 11,
  "timed_out" : false,
  "_shards" : {
    "total" : 1,
    "successful" : 1,
    "skipped" : 0,
    "failed" : 0
  },
  "hits" : {
    "total" : {
      "value" : 6,
      "relation" : "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"
        }
      },
```

(continues on the next page…)

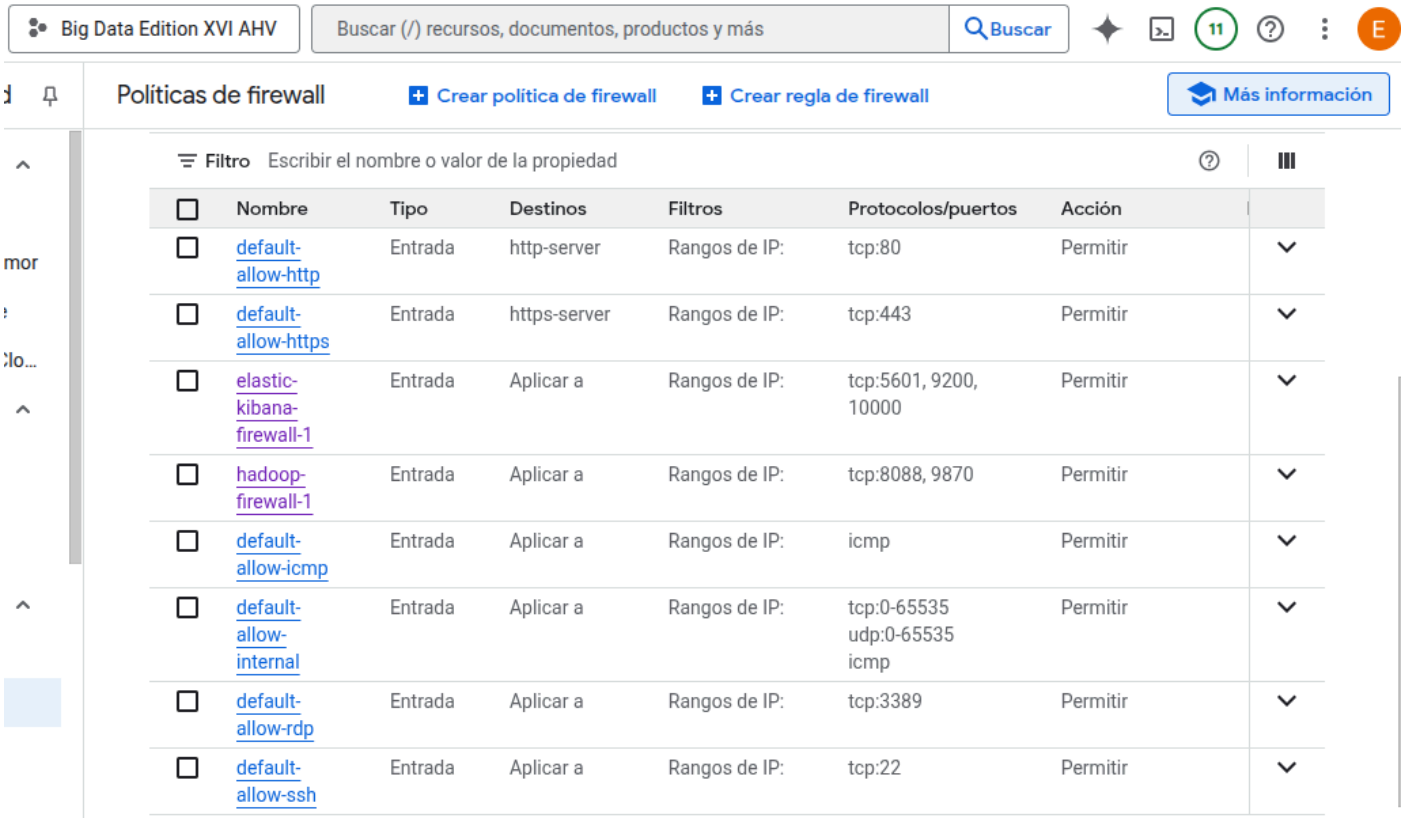
```
          "name" : "María",
          "last_name" : "López"
        }
      },
      {
        "_index" : "alumnos",
        "_id" : "5",
        "_score" : 1.0,
        "_source" : {
          "id" : 5,
          "name" : "Luis",
          "last_name" : "Martínez"
        }
      },
      {
        "_index" : "alumnos",
        "_id" : "7",
        "_score" : 1.0,
        "_source" : {
          "id" : 7,
          "name" : "Sofía",
          "last_name" : "Ramírez"
        }
      },
      {
        "_index" : "alumnos",
        "_id" : "8",
        "_score" : 1.0,
        "_source" : {
          "id" : 8,
          "name" : "Pedro",
          "last_name" : "Hernández"
        }
      }
    ]
  }
}
enairaznieh@es-hadoop-cluster-ahv-m:~$
```

PART 5 SCREENSHOTS:





Firewalls involved in this project (in magenta, the ones I set up manually):



Nombre	Tipo	Destinos	Filtros	Protocolos/puertos	Acción
default-allow-http	Entrada	http-server	Rangos de IP:	tcp:80	Permitir
default-allow-https	Entrada	https-server	Rangos de IP:	tcp:443	Permitir
elastic-kibana-firewall-1	Entrada	Aplicar a	Rangos de IP:	tcp:5601, 9200, 10000	Permitir
hadoop-firewall-1	Entrada	Aplicar a	Rangos de IP:	tcp:8088, 9870	Permitir
default-allow-icmp	Entrada	Aplicar a	Rangos de IP:	icmp	Permitir
default-allow-internal	Entrada	Aplicar a	Rangos de IP:	tcp:0-65535 udp:0-65535 icmp	Permitir
default-allow-rdp	Entrada	Aplicar a	Rangos de IP:	tcp:3389	Permitir
default-allow-ssh	Entrada	Aplicar a	Rangos de IP:	tcp:22	Permitir