

# ACTIVIDAD 2.7

Universidad de Guadalajara

Varela Tavares Gabriel

Noviembre de 2021

De la actividad 2.6 la segunda consulta **Obtener los primeros 100 registros de la tabla empleo (le puse asecg2020); *SELECT \* FROM asecg2020 LIMIT 100;*** nos mostraba cero *MAP* y cero *REDUCE*. En esta nueva actividad se pide hacer la misma consulta y ordenar por *cve\_entidad* o *cve\_subdelegación* (segunda consulta de la actividad 2.6).

## Consulta y cambios en procesos MAP y REDUCE

Se arrancan los servicios de Hadoop y de Apache Hive. Se usa la base de datos anterior y luego ***SELECT \* FROM asecg2020 ORDER BY cve\_entidad LIMIT 100;*** Esto muestra MAP:4 y REDUCE:1.

```
Terminal - oem@oem-VirtualBox: /usr/local/apache-hive
Archivo Editar Ver Terminal Pestañas Ayuda
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks determined at compile time: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1636936587774_0001, Tracking URL = http://oem-VirtualBox:8080/proxy/application_1636936587774_0001/
Kill Command = /usr/local/hadoop/bin/hadoop job -kill job_1636936587774_0001
Hadoop job information for Stage-1: number of mappers: 4; number of reducers: 1
2021-11-14 18:44:31.923 Stage-1 map = 0%, reduce = 0%
2021-11-14 18:44:37.637 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 59.12 sec
2021-11-14 18:45:03.921 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 79.86 sec
2021-11-14 18:45:10.189 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 92.3 sec
2021-11-14 18:45:16.497 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 103.24 sec
2021-11-14 18:45:19.600 Stage-1 map = 71%, reduce = 17%, Cumulative CPU 103.57 sec
2021-11-14 18:45:21.884 Stage-1 map = 86%, reduce = 17%, Cumulative CPU 111.74 sec
2021-11-14 18:45:23.878 Stage-1 map = 86%, reduce = 25%, Cumulative CPU 111.81 sec
2021-11-14 18:45:24.921 Stage-1 map = 100%, reduce = 25%, Cumulative CPU 114.14 sec
2021-11-14 18:45:25.946 Stage-1 map = 100%, reduce = 160%, Cumulative CPU 116.37 sec
MapReduce Total cumulative CPU time: 1 minutes 56 seconds 370 msec
Ended Job = job_1636936587774_0001
MapReduce Jobs Launched:
Stage-1:1: Map: 4 Reduce: 1 Cumulative CPU: 116.37 sec HDFS Read: 1116382373 HDFS Write: 9378 SUCCESS
Total MapReduce CPU Time Spent: 1 minutes 56 seconds 370 msec
OK
1 1 1 A01 0 1 101 S2 1 E6 W3 W4 1 011 0 0 0 1 1 0 0 0 317.42 317.42 0.0 0.0
1 0.0 1 1 A01 0 1 101 S2 1 E6 W2 W2 12 012 0 2 2 8 12 0 2 2 8 1788.39 0.0 296.2
1 298.34 1193.85 1 1 A01 0 1 101 S2 1 E5 W3 W4 1 010 0 1 0 1 0 0 1 0 319.41 0.0 0.0 319.41
1 0.0 1 1 A01 0 1 101 S2 1 E5 W2 W2 6 060 1 0 5 6 0 1 0 5 983.36 0.0 148.1 0.0
1 755.26 1 1 A01 0 1 101 S2 1 E4 W3 W4 1 010 0 1 0 1 0 0 1 0 319.41 0.0 0.0 319.41
1 0.0 1 1 A01 0 1 101 S2 1 E2 W2 W3 1 010 0 0 1 1 0 0 0 1 280.0 0.0 0.0 0.0
1 280.0 1 1 A01 0 1 101 S2 1 E2 W2 W2 1 010 0 0 1 1 0 0 0 1 148.1 0.0 0.0 0.0
1 148.1 1 1 A01 0 1 101 S2 1 E13 W2 W3 1 010 0 0 1 1 0 0 0 1 280.0 0.0 0.0 0.0
1 280.0 1 1 A01 0 1 101 S2 1 E13 W2 W2 1 010 0 0 1 1 0 0 0 1 148.1 0.0 0.0 0.0
1 148.1 1 1 A01 0 1 101 S2 1 E11 W2 W2 3 030 0 0 3 3 0 0 0 3 444.3 0.0 0.0 0.0
1 444.3 1 1 A01 0 1 101 S2 1 E10 W4 W6 1 010 0 1 0 1 0 0 1 0 454.98 0.0 0.0 454.98
1 0.0 1 1 A01 0 1 101 S2 1 E10 W2 W3 1 010 0 0 1 1 0 0 0 1 280.0 0.0 0.0 0.0
1 280.0 1 1 A01 0 1 101 S2 1 E10 W2 W2 14 014 0 2 0 12 14 0 2 0 12 2118.42 0.0 296.2
1 0.0 1822.22 1 1 A01 0 1 101 S1 2 E5 W2 W3 1 010 0 0 1 1 0 0 0 1 260.77 0.0 0.0 0.0
1 260.77 1 1 A01 0 1 101 S1 2 E5 W2 W2 1 010 0 0 1 1 0 0 0 1 148.1 0.0 0.0 0.0
```

También se hizo la consulta con order by *cve\_subdelegacion* de la siguiente manera: ***SELECT \* FROM asecg2020 ORDER BY cve\_subdelegacion LIMIT 100;***. Esta ejecución realiza MAP:4 y REDUCE:1.

```
Terminal - oem@oem-VirtualBox: /usr/local/apache-hive

Archivo Editar Ver Terminal Pestañas Ayuda

0.0 0.0 0.0
1 1 1
0.0 238.72 252.94 A01 0 1 101 S2 1 E5 W2 W3 2 0 2 0 0 1 1 2 0 0 1 1 491.66 0.0 0
1 19 1 Y49 9 94 9401 S4 2 E9 W3 W4 2 0 2 0 0 2 0 2 0 0 2 0 675.97 0.0 0
0.0 675.97 0.0 A01 NA 1 E1 NA NA 178 178 0 0 0 0 0 0 0 0 0.0 0.0 0
1 1 1
0.0 0.0 0.0
Time taken: 70.688 seconds, Fetched: 100 row(s)
hive> select* from asag2020 order by cve subdelegacion limit 1001
WARNING: Hive-0.12.1 is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e. spark, tez) or using Hive 1.X releases.
Query ID = oem_20211114165841_98f3f36f-e514-4e99-a437-286a9a95b4e2
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks determined at compile time: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1636936587774_0003, Tracking URL = http://oem-VirtualBox:8088/proxy/application_1636936587774_0003/
Kill Command = /usr/local/hadoop/bin/hadoop job -kill job_1636936587774_0003
Hadoop job information for Stage-1: number of mappers: 4; number of reducers: 1
2021-11-14 18:50:51.360 Stage-1 map = 0%, reduce = 0%
2021-11-14 18:51:12.415 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 45.22 sec
2021-11-14 18:51:14.470 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 50.93 sec
2021-11-14 18:51:15.511 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 62.1 sec
2021-11-14 18:51:27.015 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 84.42 sec
2021-11-14 18:51:28.002 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 89.24 sec
2021-11-14 18:51:31.172 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 91.95 sec
2021-11-14 18:51:32.196 Stage-1 map = 71%, reduce = 17%, Cumulative CPU 92.23 sec
2021-11-14 18:51:37.325 Stage-1 map = 80%, reduce = 17%, Cumulative CPU 107.25 sec
2021-11-14 18:51:38.377 Stage-1 map = 100%, reduce = 17%, Cumulative CPU 111.68 sec
2021-11-14 18:51:40.468 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 113.74 sec
MapReduce Total cumulative CPU time: 1 minutes 53 seconds 740 msec
Ended Job = job_1636936587774_0003
MapReduce Jobs Launched:
Stage-1: Map: 4 Reduce: 1 Cumulative CPU: 113.74 sec HDFS Read: 1116382405 HDFS Write: 9663 SUCCESS
Total MapReduce CPU Time Spent: 1 minutes 53 seconds 740 msec
OK
27 1 26 E64 0 1 101 S2 1 E13 W2 W2 1 0 1 0 0 0 1 1 0 0 0 1 148.1 0.0 0
0.0 148.1
27 1 26 E64 0 1 101 S2 1 E12 W2 W3 3 0 3 0 1 0 2 3 0 1 0 2 712.07 0.0 1
20.7 0.0 521.37
27 1 26 E64 0 1 101 S2 1 E12 W2 W2 5 0 5 0 1 2 2 5 0 1 2 2 740.5 0.0 1
48.1 296.2 296.2
27 1 26 E64 0 1 101 S2 1 E11 W3 W4 2 0 2 0 0 1 1 2 0 0 1 1 627.12 0.0 0
0 313.56 313.56
27 1 26 E64 0 1 101 S2 1 E11 W2 W3 1 0 1 0 1 0 0 1 0 1 0 0 200.0 0.0 2
80.0 0.0 0.0
27 1 26 E64 0 1 101 S2 1 E10 W2 W3 4 0 4 0 0 2 2 4 0 0 2 2 909.69 0.0 0
0 427.07 482.62
27 1 26 E64 0 1 101 S2 1 E10 W2 W2 14 0 14 0 0 2 12 14 0 0 2 12 2135.92 0.0 0
0 323.44 1812.48
27 1 26 E64 0 1 101 S2 1 E10 W11 W17 1 0 1 0 0 1 0 1 0 0 1 0 1488.47 0.0 0
0 1488.47 0.0
27 1 26 E64 0 1 101 S1 2 E0 W0 W0 1 0 1 0 0 1 0 1 0 0 1 0 360.0 0.0 0
```