



INFSCI 2750: Cloud Computing

mini project 2

Haotian Wu(haw145); Pinhao Wang(piw17); Zhecheng Qi(zhq27)

March 10, 2023

Objective:

The objective of this mini project is to learn and develop applications using Apache Cassandra. You will use the previously assigned VMs.

Part 1: Setting up Cassandra: (50 points)

```
ubuntu@ccprojects-22:~$ nodetool status
Datacenter: datacenter1
=====
Status=Up/Down
// State=Normal/Leaving/Joining/Moving
-- Address            Load            Tokens          Owns            Host ID                               Rack
UN  10.254.3.110       65.36 MiB       256             ?              700facc1-d8ce-4df3-94c9-7b86818a9a31 rack1
UN  10.254.1.29        65.52 MiB       256             ?              e26ee5b9-7f0f-4ebf-aa9a-2d98f0f0129c rack1
UN  10.254.1.131       65.51 MiB       256             ?              7d6e1dce-7856-49c7-b541-0283983d2c23 rack1

Note: Non-system keyspaces don't have the same replication settings, effective ownership information
ubuntu@ccprojects-22:~$ cqlsh CC-demo-1
Connected to Test Cluster at CC-demo-1:9042.
[cqlsh 5.0.1 | Cassandra 3.11.13 | CQL spec 3.4.4 | Native protocol v4]
Use HELP for help.
cqlsh> exit
ubuntu@ccprojects-22:~$ cqlsh CC-demo-2
Connected to Test Cluster at CC-demo-2:9042.
[cqlsh 5.0.1 | Cassandra 3.11.13 | CQL spec 3.4.4 | Native protocol v4]
Use HELP for help.
cqlsh> exit
ubuntu@ccprojects-22:~$ cqlsh CC-demo-3
Connected to Test Cluster at CC-demo-3:9042.
[cqlsh 5.0.1 | Cassandra 3.11.13 | CQL spec 3.4.4 | Native protocol v4]
Use HELP for help.
cqlsh> exit
ubuntu@ccprojects-22:~$
```

Figure 1: Configure a Cassandra distribution in our cluster of VMs.

Part 2: Import Data into Cassandra (25 points)

```
cqlsh:test4> select * from access;
```

id	date	ip	path	status
4317	[23/Aug/2009:13:18:28 -0700]	10.211.47.159	GET /assets/img/closetlabel.gif HTTP/1.1	-304
3372	[13/Aug/2009:15:15:24 -0700]	10.82.30.199	GET /favicon.ico HTTP/1.1	404 209
1584	[10/Aug/2009:19:57:36 -0700]	10.211.47.159	GET /assets/img/home-logo.png HTTP/1.1	-304
7034	[03/Sep/2009:22:47:50 -0700]	10.211.47.159	GET /assets/css/960.css HTTP/1.1	-304
9892	[15/Oct/2009:21:33:15 -0700]	10.211.47.159	GET /assets/js/the-associates.js HTTP/1.1	-304
9640	[15/Oct/2009:21:21:57 -0700]	10.211.47.159	GET /displaytitle.php?id=10 HTTP/1.1	200 11330
9067	[15/Oct/2009:20:31:43 -0700]	10.211.47.159	GET /assets/js/jquery.fancybox.min.js HTTP/1.1	-304
4830	[26/Aug/2009:12:25:24 -0700]	10.211.47.159	GET /displaytitle.php?id=12 HTTP/1.1	200 8849
2731	[11/Aug/2009:20:34:44 -0700]	10.211.47.159	GET /assets/img/dummy/primary-news-2.jpg HTTP/1.1	-304
5056	[01/Sep/2009:02:24:19 -0700]	10.216.113.172	GET /assets/css/960.css HTTP/1.1	-304
6428	[03/Sep/2009:22:14:53 -0700]	10.211.47.159	GET /assets/js/the-associates.js HTTP/1.1	-304
2713	[11/Aug/2009:20:34:40 -0700]	10.211.47.159	GET /assets/img/search-button.gif HTTP/1.1	-304
769	[29/Jul/2009:09:12:19 -0700]	10.153.239.5	GET /assets/img/home-logo.png HTTP/1.1	-304
9973	[15/Oct/2009:21:35:37 -0700]	10.211.47.159	GET /images/clientlogos/0000/0003/ HTTP/1.1	404 227
1863	[10/Aug/2009:20:43:41 -0700]	10.211.47.159	GET /release-schedule/ HTTP/1.1	200 9141
9495	[15/Oct/2009:21:05:56 -0700]	10.211.47.159	GET /images/filmpics/0000/0051/quietman1.jpeg HTTP/1.1	-304
7863	[08/Oct/2009:20:29:24 -0700]	10.211.47.159	GET /assets/css/960.css HTTP/1.1	-304
2453	[11/Aug/2009:12:24:28 -0700]	10.211.47.159	GET /assets/js/lightbox.js HTTP/1.1	-304
2302	[11/Aug/2009:12:09:07 -0700]	10.211.47.159	GET / HTTP/1.1	200 7616
4120	[22/Aug/2009:12:56:29 -0700]	10.211.47.159	GET /assets/img/loading.gif HTTP/1.1	200 2767
5801	[02/Sep/2009:08:35:46 -0700]	10.216.113.172	GET /assets/js/lowpro.js HTTP/1.1	-304
9979	[15/Oct/2009:21:35:37 -0700]	10.211.47.159	GET /images/filmpics/0000/0051/quietman1_thumb.jpeg HTTP/1.1	-304
7540	[07/Sep/2009:15:29:53 -0700]	10.82.30.199	GET / HTTP/1.1	200 6700
3733	[18/Aug/2009:18:03:01 -0700]	10.211.47.159	GET /assets/css/960.css HTTP/1.1	-304
1765	[10/Aug/2009:20:23:58 -0700]	10.211.47.159	GET /assets/img/dummy/films/thumbs/mum-and-dad.jpg HTTP/1.1	-304
7789	[07/Oct/2009:19:32:48 -0700]	10.211.47.159	GET /images/clientlogos/0000/0003/ HTTP/1.1	404 227
9527	[15/Oct/2009:21:11:05 -0700]	10.211.47.159	GET /assets/js/the-associates.js HTTP/1.1	-304
4822	[26/Aug/2009:12:23:12 -0700]	10.211.47.159	GET /assets/img/dummy/home-media-block/mum-and-dad.jpg HTTP/1.1	-304
4650	[24/Aug/2009:13:18:41 -0700]	10.211.47.159	GET /assets/css/the-associates.css HTTP/1.1	-304
1580	[10/Aug/2009:19:57:35 -0700]	10.211.47.159	GET /assets/js/lowpro.js HTTP/1.1	-304
7018	[03/Sep/2009:22:47:36 -0700]	10.211.47.159	GET /images/filmpics/0000/0053/quietman2_thumb.jpeg HTTP/1.1	-304
7181	[04/Sep/2009:02:58:29 -0700]	10.216.113.172	GET /displaytitle.php?id=12 HTTP/1.1	200 10973
7539	[07/Sep/2009:04:16:02 -0700]	10.216.113.172	GET /images/filmpics/0000/0067/beerleague2.jpeg HTTP/1.1	200 2840827
2062	[10/Aug/2009:20:55:12 -0700]	10.211.47.159	GET /assets/js/the-associates.js HTTP/1.1	-304
4701	[24/Aug/2009:15:35:57 -0700]	10.211.47.159	GET /images/filmpics/0000/0077/mike_thumb.jpg HTTP/1.1	-304
1199	[30/Jul/2009:02:34:58 -0700]	10.216.113.172	GET /assets/img/dummy/secondary-news-1.jpg HTTP/1.1	200 5766
3939	[21/Aug/2009:04:17:10 -0700]	10.216.113.172	GET /trailers/ HTTP/1.1	200 11099
23	[15/Jul/2009:15:50:51 -0700]	10.223.157.186	GET /assets/css/960.css HTTP/1.1	-304
7099	[03/Sep/2009:22:52:17 -0700]	10.211.47.159	GET /images/filmpics/0000/0057/quietman4.jpeg HTTP/1.1	-304
5900	[02/Sep/2009:14:56:36 -0700]	10.211.47.159	GET /assets/img/closetlabel.gif HTTP/1.1	-304
114	[15/Jul/2009:21:24:17 -0700]	10.223.157.186	GET /assets/img/dummy/film-page-media.jpg HTTP/1.1	200 110997
1548	[10/Aug/2009:19:56:54 -0700]	10.211.47.159	GET /assets/img/loading.gif HTTP/1.1	-304
9589	[15/Oct/2009:21:16:47 -0700]	10.211.47.159	GET /displaytitle.php?id=10 HTTP/1.1	200 11028

Figure 2: Imported data and data table structure

Part 3: Operate Data in Cassandra (25 points)

Problems1: How many hits were made to the website item “/assets/img/release-schedule-logo.png”?

```
cqlsh:test4> SELECT COUNT (*) AS path_count FROM access where path = 'GET /assets/img/release-schedule-logo.png HTTP/1.1' ALLOW FILTERING;
```

path_count
75

Figure 3: Hits of '/assets/img/release-schedule-logo.png'

Problems2: How many hits were made from the IP: 10.207.188.188?

```
cqlsh:test4> select count(*) as ip_count from access where ip = '10.207.188.188' ALLOW FILTERING;
```

ip_count
2

Figure 4: Hits of 'IP: 10.207.188.188'

Problems3: Which path in the website has been hit most? How many hits were made to the path?

```
... INFO [0],
cqlsh:test4> select groupcountmaxii(path) from access;

test4.groupcountmaxii(path)
-----
{'GET /assets/css/the-associates.css HTTP/1.1': 488}

(1 rows)

Warnings :
Aggregation query used without partition key
```

Figure 5: path in the website has been hit most

Problems4: Which IP accesses the website most? How many accesses were made by it?

```
cqlsh:test4> select groupcountmaxii(ip) from access;

test4.groupcountmaxii(ip)
-----
{'10.211.47.159': 6687}

(1 rows)

Warnings :
Aggregation query used without partition key
```

Figure 6: IP accesses the website most