



télécom  
saint-étienne

école d'ingénieurs / nouvelles technologies

## NoSQL : Cassandra

Par : KAROUL Ayoub, GE Li

On commence par créer un keyspace 'meteo' qui sera notre conteneur de données en précisant le facteur de répllication ainsi que la stratégie de distribution des réplicas.

Afin d'assurer la fiabilité et la tolérance aux pannes, on va stocker nos données sur 3 nœuds différents du cluster.

```
root@karoul-VirtualBox: /home/karoul
cqlsh> CREATE KEYSPACE meteo WITH REPLICATION = { 'class' : 'SimpleStrategy',
'replication_factor' : 3 };
cqlsh>
```

On crée une table 'stationmeteo' qui va modéliser l'ensemble des données remontées par les stations météo :

```
cqlsh>
cqlsh> use meteo;
cqlsh:meteo>
cqlsh:meteo> CREATE TABLE stationmeteo(
... id_station text,
... longitude text,
... latitude text,
... horodate timestamp,
... temperature text,
... humidite text,
... primary key (id_station, horodate));
cqlsh:meteo>
```

Sa clé primaire est la paire de colonnes : (id\_station et horodate) tel que :

« id\_station » : clé de partition de type 'text'.

« horodate » : clé de clustering de type 'timestamp'.

- Une station météo possède un identifiant unique « id\_station » et ne peut pas fournir plusieurs mesures au même « horodate ».
- Plusieurs stations météo peuvent fournir des mesures au même « horodate ».

Pour insérer les données dans notre table on va utiliser la commande suivante :

```
root@karoul-VirtualBox: /home/karoul
cqlsh:meteo> INSERT INTO stationmeteo (id_station, longitude, latitude, horodate, temperature,
humidite) VALUES ('SE001', '3.45', '50.2145', '2020-02-20T01:45:16+0000', '10°C', '40%');
cqlsh:meteo>
cqlsh:meteo> select * from stationmeteo
... ;

id_station | horodate | humidite | latitude | longitude | temperature
-----+-----+-----+-----+-----+-----
SE001 | 2020-02-20 01:45:16.000000+0000 | 40% | 50.2145 | 3.45 | 10°C

(1 rows)
cqlsh:meteo>
```

Dans l'exemple précédant on a ajouté une station 'SE001' et ses mesures (température, humidité... etc.) à la date 'horodate'.

- De la même façon, on va ajouter les mesures des stations 'SE002' et 'SE003' au même 'horodate'.

```
INSERT INTO stationmeteo (id_station, longitude, latitude, horodate, temperature, humidite)
VALUES ('SE002', '3.01', '42.2346', '2020-02-20T01:45:16+0000', '11°C', '42%');
```

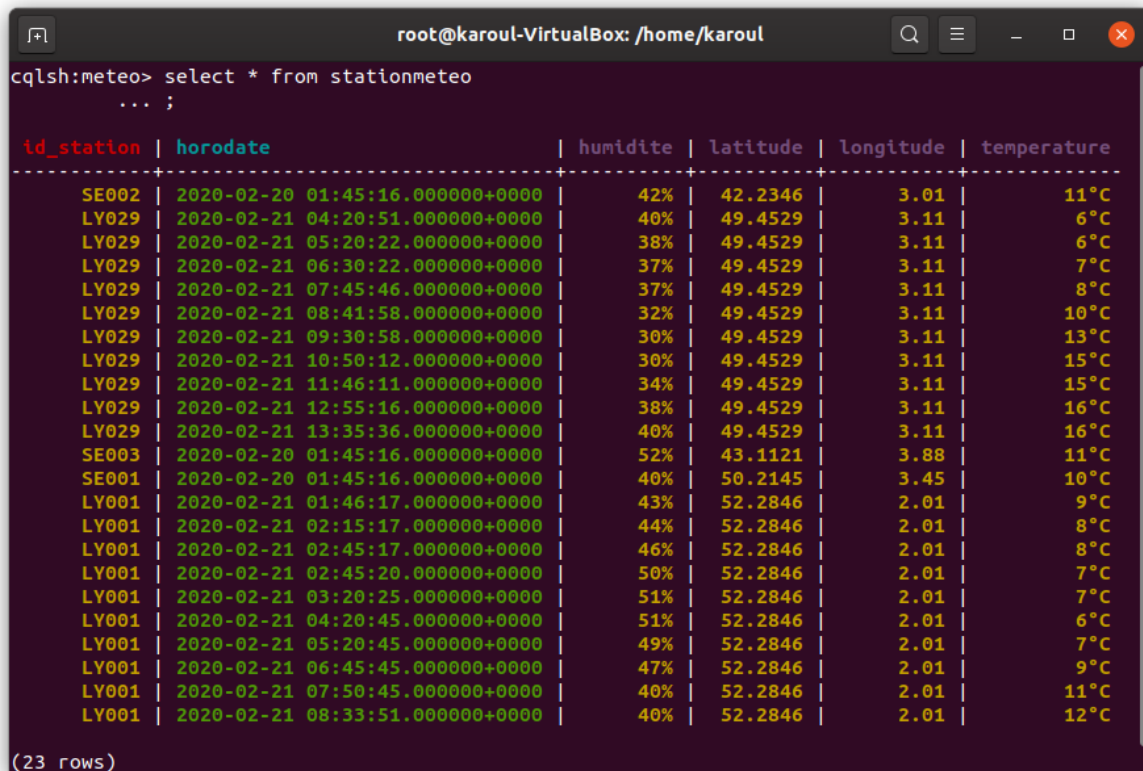
```
INSERT INTO stationmeteo (id_station, longitude, latitude, horodate, temperature, humidite)
VALUES ('SE003', '3.88', '43.1121', '2020-02-20T01:45:16+0000', '11°C', '52%');
```

- Maintenant, on va ajouter les mesures des stations 'LY001' et 'LY029' pour des 'horodate' différents :

```
cqlsh:meteo> INSERT INTO stationmeteo (id_station, longitude, latitude, horodate, temperature,
humidite) VALUES ('LY001', '2.01', '52.2846', '2020-02-21T02:45:20+0000', '7°C', '50%');
cqlsh:meteo> INSERT INTO stationmeteo (id_station, longitude, latitude, horodate, temperature,
humidite) VALUES ('LY001', '2.01', '52.2846', '2020-02-21T03:20:25+0000', '7°C', '51%');
cqlsh:meteo> INSERT INTO stationmeteo (id_station, longitude, latitude, horodate, temperature,
humidite) VALUES ('LY001', '2.01', '52.2846', '2020-02-21T04:20:45+0000', '6°C', '51%');
cqlsh:meteo> INSERT INTO stationmeteo (id_station, longitude, latitude, horodate, temperature,
humidite) VALUES ('LY001', '2.01', '52.2846', '2020-02-21T05:20:45+0000', '7°C', '49%');
cqlsh:meteo> INSERT INTO stationmeteo (id_station, longitude, latitude, horodate, temperature,
humidite) VALUES ('LY001', '2.01', '52.2846', '2020-02-21T06:45:45+0000', '9°C', '47%');
cqlsh:meteo> INSERT INTO stationmeteo (id_station, longitude, latitude, horodate, temperature,
humidite) VALUES ('LY001', '2.01', '52.2846', '2020-02-21T07:50:45+0000', '11°C', '40%');
cqlsh:meteo> INSERT INTO stationmeteo (id_station, longitude, latitude, horodate, temperature,
humidite) VALUES ('LY001', '2.01', '52.2846', '2020-02-21T08:33:51+0000', '12°C', '40%');
cqlsh:meteo>
cqlsh:meteo>
```

```
cqlsh:meteo> INSERT INTO stationmeteo (id_station, longitude, latitude, horodate, temperature,
humidite) VALUES ('LY029', '3.11', '49.4529', '2020-02-21T05:20:22+0000', '6°C', '38%');
cqlsh:meteo> INSERT INTO stationmeteo (id_station, longitude, latitude, horodate, temperature,
humidite) VALUES ('LY029', '3.11', '49.4529', '2020-02-21T06:30:22+0000', '7°C', '37%');
cqlsh:meteo> INSERT INTO stationmeteo (id_station, longitude, latitude, horodate, temperature,
humidite) VALUES ('LY029', '3.11', '49.4529', '2020-02-21T07:45:46+0000', '8°C', '37%');
cqlsh:meteo> INSERT INTO stationmeteo (id_station, longitude, latitude, horodate, temperature,
humidite) VALUES ('LY029', '3.11', '49.4529', '2020-02-21T08:41:58+0000', '10°C', '32%');
cqlsh:meteo> INSERT INTO stationmeteo (id_station, longitude, latitude, horodate, temperature,
humidite) VALUES ('LY029', '3.11', '49.4529', '2020-02-21T09:30:58+0000', '13°C', '30%');
cqlsh:meteo> INSERT INTO stationmeteo (id_station, longitude, latitude, horodate, temperature,
humidite) VALUES ('LY029', '3.11', '49.4529', '2020-02-21T10:50:12+0000', '15°C', '30%');
cqlsh:meteo> INSERT INTO stationmeteo (id_station, longitude, latitude, horodate, temperature,
humidite) VALUES ('LY029', '3.11', '49.4529', '2020-02-21T11:46:11+0000', '15°C', '34%');
cqlsh:meteo> INSERT INTO stationmeteo (id_station, longitude, latitude, horodate, temperature,
humidite) VALUES ('LY029', '3.11', '49.4529', '2020-02-21T12:55:16+0000', '16°C', '38%');
cqlsh:meteo> INSERT INTO stationmeteo (id_station, longitude, latitude, horodate, temperature,
humidite) VALUES ('LY029', '3.11', '49.4529', '2020-02-21T13:35:36+0000', '16°C', '40%');
cqlsh:meteo>
```

On peut visualiser notre table pour vérifier l'insertion des mesures :



root@karoul-VirtualBox: /home/karoul

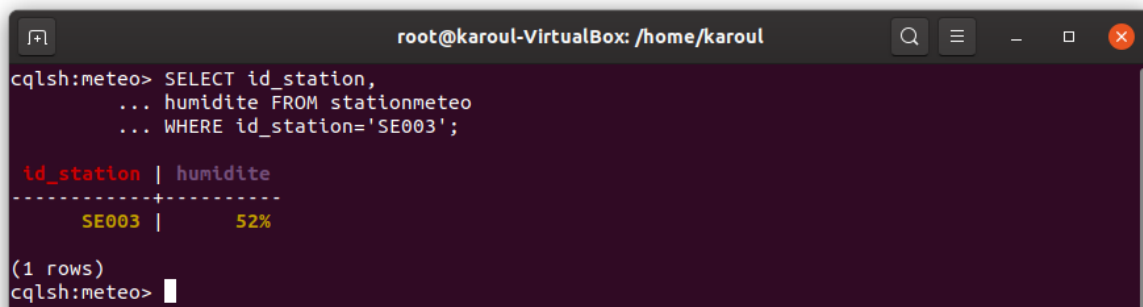
```
cqlsh:meteo> select * from stationmeteo
... ;
```

id_station	horodate	humidite	latitude	longitude	temperature
SE002	2020-02-20 01:45:16.000000+0000	42%	42.2346	3.01	11°C
LY029	2020-02-21 04:20:51.000000+0000	40%	49.4529	3.11	6°C
LY029	2020-02-21 05:20:22.000000+0000	38%	49.4529	3.11	6°C
LY029	2020-02-21 06:30:22.000000+0000	37%	49.4529	3.11	7°C
LY029	2020-02-21 07:45:46.000000+0000	37%	49.4529	3.11	8°C
LY029	2020-02-21 08:41:58.000000+0000	32%	49.4529	3.11	10°C
LY029	2020-02-21 09:30:58.000000+0000	30%	49.4529	3.11	13°C
LY029	2020-02-21 10:50:12.000000+0000	30%	49.4529	3.11	15°C
LY029	2020-02-21 11:46:11.000000+0000	34%	49.4529	3.11	15°C
LY029	2020-02-21 12:55:16.000000+0000	38%	49.4529	3.11	16°C
LY029	2020-02-21 13:35:36.000000+0000	40%	49.4529	3.11	16°C
SE003	2020-02-20 01:45:16.000000+0000	52%	43.1121	3.88	11°C
SE001	2020-02-20 01:45:16.000000+0000	40%	50.2145	3.45	10°C
LY001	2020-02-21 01:46:17.000000+0000	43%	52.2846	2.01	9°C
LY001	2020-02-21 02:15:17.000000+0000	44%	52.2846	2.01	8°C
LY001	2020-02-21 02:45:17.000000+0000	46%	52.2846	2.01	8°C
LY001	2020-02-21 02:45:20.000000+0000	50%	52.2846	2.01	7°C
LY001	2020-02-21 03:20:25.000000+0000	51%	52.2846	2.01	7°C
LY001	2020-02-21 04:20:45.000000+0000	51%	52.2846	2.01	6°C
LY001	2020-02-21 05:20:45.000000+0000	49%	52.2846	2.01	7°C
LY001	2020-02-21 06:45:45.000000+0000	47%	52.2846	2.01	9°C
LY001	2020-02-21 07:50:45.000000+0000	40%	52.2846	2.01	11°C
LY001	2020-02-21 08:33:51.000000+0000	40%	52.2846	2.01	12°C

(23 rows)

Il est possible maintenant de faire des requêtes sur les données stockées.

- Par exemple pour avoir uniquement le taux d'humidité de la station 'SE003' :



root@karoul-VirtualBox: /home/karoul

```
cqlsh:meteo> SELECT id_station,
... humidite FROM stationmeteo
... WHERE id_station='SE003';
```

id_station	humidite
SE003	52%

(1 rows)

```
cqlsh:meteo>
```

- Pour afficher toutes les mesures de la station 'LY001', on précise son 'id\_station' dans la requête comme suit :

```

root@karoul-VirtualBox: /home/karoul
cqlsh:meteo>
cqlsh:meteo> SELECT * FROM stationmeteo WHERE id_station='LY001';

 id_station | horodate | humidite | latitude | longitude | temperature
-----|-----|-----|-----|-----|-----
 LY001 | 2020-02-21 01:46:17.000000+0000 | 43% | 52.2846 | 2.01 | 9°C
 LY001 | 2020-02-21 02:15:17.000000+0000 | 44% | 52.2846 | 2.01 | 8°C
 LY001 | 2020-02-21 02:45:17.000000+0000 | 46% | 52.2846 | 2.01 | 8°C
 LY001 | 2020-02-21 02:45:20.000000+0000 | 50% | 52.2846 | 2.01 | 7°C
 LY001 | 2020-02-21 03:20:25.000000+0000 | 51% | 52.2846 | 2.01 | 7°C
 LY001 | 2020-02-21 04:20:45.000000+0000 | 51% | 52.2846 | 2.01 | 6°C
 LY001 | 2020-02-21 05:20:45.000000+0000 | 49% | 52.2846 | 2.01 | 7°C
 LY001 | 2020-02-21 06:45:45.000000+0000 | 47% | 52.2846 | 2.01 | 9°C
 LY001 | 2020-02-21 07:50:45.000000+0000 | 40% | 52.2846 | 2.01 | 11°C
 LY001 | 2020-02-21 08:33:51.000000+0000 | 40% | 52.2846 | 2.01 | 12°C

(10 rows)
cqlsh:meteo>

```

- Pour afficher toutes les mesures de la station 'LY029' le 21-02-2020 entre 06h00 et 12h00 :

```

cqlsh:meteo> SELECT * FROM stationmeteo WHERE id_station = 'LY029' AND horodate >= '2020-02-21T
06:00:00+0000' AND horodate <= '2020-02-21T12:00:00+0000';

 id_station | horodate | humidite | latitude | longitude | temperature
-----|-----|-----|-----|-----|-----
 LY029 | 2020-02-21 06:30:22.000000+0000 | 37% | 49.4529 | 3.11 | 7°C
 LY029 | 2020-02-21 07:45:46.000000+0000 | 37% | 49.4529 | 3.11 | 8°C
 LY029 | 2020-02-21 08:41:58.000000+0000 | 32% | 49.4529 | 3.11 | 10°C
 LY029 | 2020-02-21 09:30:58.000000+0000 | 30% | 49.4529 | 3.11 | 13°C
 LY029 | 2020-02-21 10:50:12.000000+0000 | 30% | 49.4529 | 3.11 | 15°C
 LY029 | 2020-02-21 11:46:11.000000+0000 | 34% | 49.4529 | 3.11 | 15°C

(6 rows)
cqlsh:meteo>

```