Assignment 6 - Introduction to ETCD

For any queries please email - tacloudcomputing2024@gmail.com or ectacloudcomputing2024@gmail.com

Deliverables for this assignment -

1(a) - Successful installation of etcd
2(a) - Etcd logs after starting
2(b) - get/put SRN
3(a) - All three nodes up and running in the cluster
4(a) - Output of etcdctl endpoint status before leader termination
4(b) - Output of etcdctl endpoint status after leader termination

A) Introduction to etcd

What is etcd?

Etcd is a distributed key-value store designed to reliably store data across a cluster of machines. It is widely used in distributed systems and cloud-native environments as a core component for storing configuration data, coordinating distributed applications, and implementing distributed locks and leader election. Etcd was developed by CoreOS, a company known for its focus on building scalable infrastructure solutions.

From the Cloud Computing Course point of view, etcd is an alternative to Apache Zookeeper, which some applications are moving away from these days (Eg: Kafka no longer supports Zookeeper; instead, they have their own alternative called KRaft).

Use Cases of etcd:

1. **Service Discovery:** Etcd is commonly used for service discovery in microservices architectures, allowing services to dynamically register and discover each other's endpoints.
2. **Distributed Locks:** Etcd provides distributed locking primitives that enable coordination between distributed processes, ensuring mutual exclusion and preventing race conditions.
3. **Leader Election:** Etcd implements the Raft consensus algorithm, allowing applications to perform leader election to coordinate distributed tasks and ensure high availability.

Some notable companies and applications that use etcd include:

1. **Kubernetes**: Etcd is a critical component of Kubernetes, a popular container orchestration platform. It serves as the primary datastore for storing

configuration data, cluster state, and other critical information required for Kubernetes to manage containerized applications effectively.

2. **Netflix**: Netflix, the streaming giant, relies on etcd to manage the configuration settings and service discovery in its distributed microservices architecture. Etcd helps ensure consistency and reliability across Netflix's vast infrastructure.
3. **Uber**: Uber leverages etcd for service discovery, dynamic configuration management, and coordination across its distributed architecture. Etcd helps Uber maintain reliability and scalability as it handles millions of transactions daily.

B) Installation

Note: It is recommended to carry out this assignment on Linux or Unix-based distributions, as Etcd doesn't have official installation packages for Windows.

On Linux, the commands `sudo apt install etcd` and `sudo apt install etcdctl` will do the job, but these will install older versions by default. To install the latest stable version, copy and paste the commands below into your terminal:

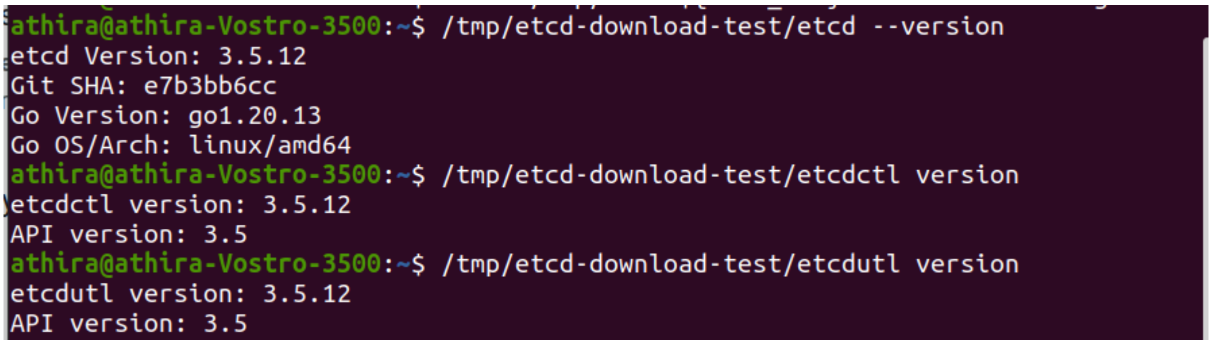
Linux:

```
ETCD_VER=v3.5.12

GITHUB_URL=https://github.com/etcd-io/etcd/releases/download
DOWNLOAD_URL=${GITHUB_URL}

rm -f /tmp/etcd-${ETCD_VER}-linux-amd64.tar.gz
rm -rf /tmp/etcd-download-test && mkdir -p /tmp/etcd-download-test

curl -L ${DOWNLOAD_URL}/${ETCD_VER}/etcd-${ETCD_VER}-linux-amd64.tar.gz -o
/tmp/etcd-${ETCD_VER}-linux-amd64.tar.gz

tar xzvf /tmp/etcd-${ETCD_VER}-linux-amd64.tar.gz -C /tmp/etcd-download-test
--strip-components=1

rm -f /tmp/etcd-${ETCD_VER}-linux-amd64.tar.gz

/tmp/etcd-download-test/etcd --version
/tmp/etcd-download-test/etcdctl version
/tmp/etcd-download-test/etcdutl version
```

MacOS (Darwin):

```
ETCD_VER=v3.5.12

GITHUB_URL=https://github.com/etcd-io/etcd/releases/download
```

```
DOWNLOAD_URL=${GITHUB_URL}

rm -f /tmp/etcd-${ETCD_VER}-darwin-amd64.zip
rm -rf /tmp/etcd-download-test && mkdir -p /tmp/etcd-download-test

curl -L ${DOWNLOAD_URL}/${ETCD_VER}/etcd-${ETCD_VER}-darwin-amd64.zip -o
/tmp/etcd-${ETCD_VER}-darwin-amd64.zip
unzip /tmp/etcd-${ETCD_VER}-darwin-amd64.zip -d /tmp && rm -f
/tmp/etcd-${ETCD_VER}-darwin-amd64.zip
mv /tmp/etcd-${ETCD_VER}-darwin-amd64/* /tmp/etcd-download-test && rm -rf mv
/tmp/etcd-${ETCD_VER}-darwin-amd64

/tmp/etcd-download-test/etcd --version
/tmp/etcd-download-test/etcdctl version
/tmp/etcd-download-test/etcdutl version
```

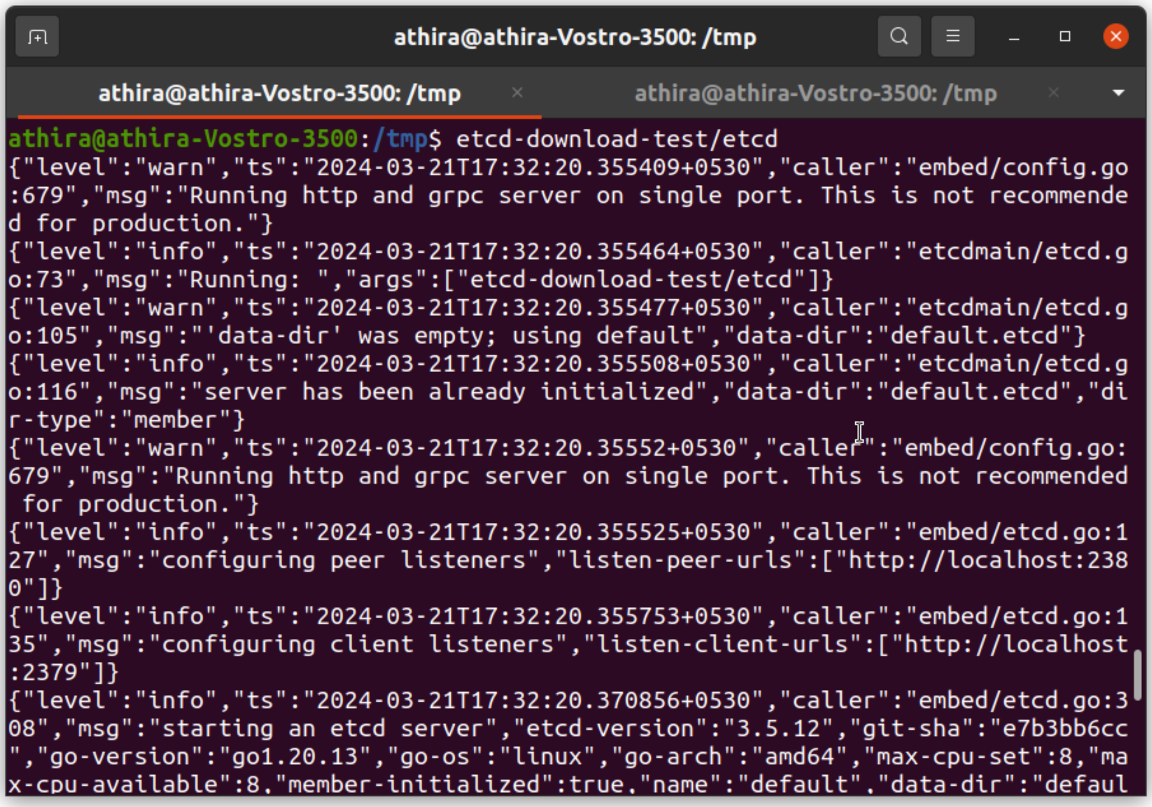


```
athira@athira-Vostro-3500:~$ /tmp/etcd-download-test/etcd --version
etcd Version: 3.5.12
Git SHA: e7b3bb6cc
Go Version: go1.20.13
Go OS/Arch: linux/amd64
athira@athira-Vostro-3500:~$ /tmp/etcd-download-test/etcdctl version
etcdctl version: 3.5.12
API version: 3.5
athira@athira-Vostro-3500:~$ /tmp/etcd-download-test/etcdutl version
etcdutl version: 3.5.12
API version: 3.5
```

Screenshot 1(a) - Successful installation of etcd

C) ETCD as a key-value store

We will first start etcd by simply using the command “etcd”. Since we want to use the version we downloaded, in a terminal window, start etcd using the command `etcd-download-test/etcd` as shown below:



```
athira@athira-Vostro-3500:/tmp$ etcd-download-test/etcd
{"level":"warn","ts":"2024-03-21T17:32:20.355409+0530","caller":"embed/config.go
:679","msg":"Running http and grpc server on single port. This is not recommende
d for production."}
{"level":"info","ts":"2024-03-21T17:32:20.355464+0530","caller":"etcdmain/etcd.g
o:73","msg":"Running: ","args":["etcd-download-test/etcd"]}
{"level":"warn","ts":"2024-03-21T17:32:20.355477+0530","caller":"etcdmain/etcd.g
o:105","msg":"'data-dir' was empty; using default","data-dir":"default.etcd"}
{"level":"info","ts":"2024-03-21T17:32:20.355508+0530","caller":"etcdmain/etcd.g
o:116","msg":"server has been already initialized","data-dir":"default.etcd","di
r-type":"member"}
{"level":"warn","ts":"2024-03-21T17:32:20.35552+0530","caller":"embed/config.go:
679","msg":"Running http and grpc server on single port. This is not recommended
 for production."}
{"level":"info","ts":"2024-03-21T17:32:20.355525+0530","caller":"embed/etcd.go:1
27","msg":"configuring peer listeners","listen-peer-urls":["http://localhost:238
0"]}
{"level":"info","ts":"2024-03-21T17:32:20.355753+0530","caller":"embed/etcd.go:1
35","msg":"configuring client listeners","listen-client-urls":["http://localhost
:2379"]}
{"level":"info","ts":"2024-03-21T17:32:20.370856+0530","caller":"embed/etcd.go:3
08","msg":"starting an etcd server","etcd-version":"3.5.12","git-sha":"e7b3bb6cc
","go-version":"go1.20.13","go-os":"linux","go-arch":"amd64","max-cpu-set":8,"ma
x-cpu-available":8,"member-initialized":true,"name":"default","data-dir":"defaul
```

Screenshot 2(a)

In a new terminal window, we will create a key value pair to store data in the node we currently have. This is done using -
`etcdctl --endpoints=localhost:2379 put srn <your_srn>`

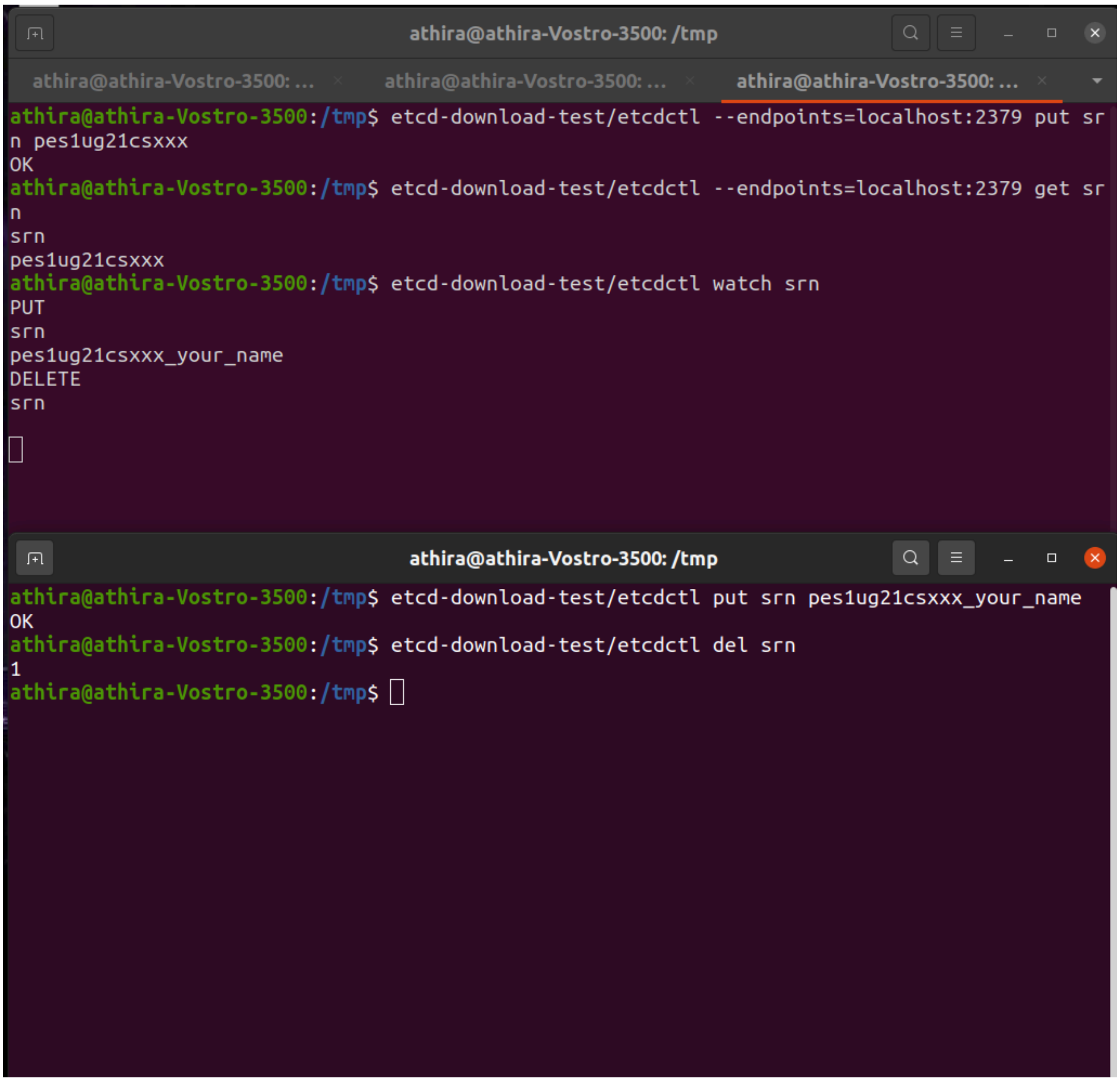
The output should simply say OK.
Similarly, `etcdctl --endpoints=localhost:2379 get srn` prints the key and the value you have created.

To watch for changes in the data, type in `etcdctl watch srn`
In a new terminal window, we will change the value, and later delete the SRN key we just created. This is done using

1) `etcdctl --endpoints=localhost:2379 put srn <your_srn_you_name>`
2) `etcd del srn`

When you do this, you will be able to see the status of your key in the output of the watch command. Take a screenshot of both terminals together, as shown below.

Screenshot 2(b)


athira@athira-Vostro-3500: /tmp
athira@athira-Vostro-3500: ...
athira@athira-Vostro-3500: ...
athira@athira-Vostro-3500: ...

```
athira@athira-Vostro-3500:/tmp$ etcd-download-test/etcdctl --endpoints=localhost:2379 put sr
n pes1ug21csxxx
OK
athira@athira-Vostro-3500:/tmp$ etcd-download-test/etcdctl --endpoints=localhost:2379 get sr
n
srn
pes1ug21csxxx
athira@athira-Vostro-3500:/tmp$ etcd-download-test/etcdctl watch srn
PUT
srn
pes1ug21csxxx_your_name
DELETE
srn
```

athira@athira-Vostro-3500: /tmp

```
athira@athira-Vostro-3500:/tmp$ etcd-download-test/etcdctl put srn pes1ug21csxxx_your_name
OK
athira@athira-Vostro-3500:/tmp$ etcd-download-test/etcdctl del srn
1
athira@athira-Vostro-3500:/tmp$ 
```

Note: In any of the above steps, if you get a connection refused error, 1) ensure you have started etcd, or 2) try changing the port numbers to something else

D) Create a cluster

First, we will set up three nodes to create our cluster. To make things easier, we won't have different machines or VMs, but just three separate terminals with different port numbers for each node.
We will create three nodes named etcd-1, etcd-2 and etcd-3. The configuration files are given below:

etcd1.conf.yml

```
name: etcd-1
data-dir: /etcd1
initial-advertise-peer-urls: http://127.0.0.2:2390
listen-peer-urls: http://127.0.0.2:2390
listen-client-urls: http://127.0.0.2:2369,http://localhost:2369
advertise-client-urls: http://127.0.0.2:2369
initial-cluster-token: etcd-cluster
initial-cluster:
etcd-1=http://127.0.0.2:2390,etcd-2=http://127.0.0.2:2391,etcd-3=h
ttp://127.0.0.2:2392
initial-cluster-state: new
```

etcd2.conf.yml

```
name: etcd-2
data-dir: /etcd2
initial-advertise-peer-urls: http://127.0.0.2:2391
listen-peer-urls: http://127.0.0.2:2391
listen-client-urls: http://127.0.0.2:2359,http://localhost:2359
advertise-client-urls: http://127.0.0.2:2359
initial-cluster-token: etcd-cluster
initial-cluster:
etcd-1=http://127.0.0.2:2390,etcd-2=http://127.0.0.2:2391,etcd-3=h
ttp://127.0.0.2:2392
initial-cluster-state: new
```

etcd3.conf.yml

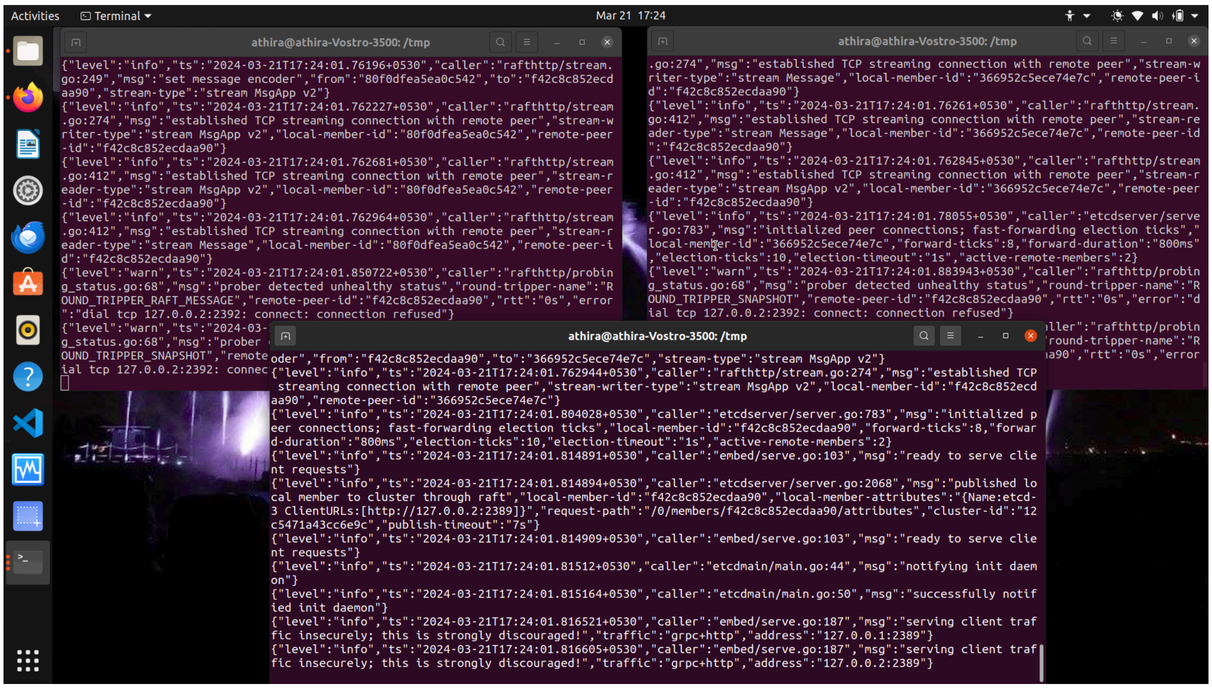
```
name: etcd-3
data-dir: /etcd3
initial-advertise-peer-urls: http://127.0.0.2:2392
listen-peer-urls: http://127.0.0.2:2392
listen-client-urls: http://127.0.0.2:2389,http://localhost:2389
advertise-client-urls: http://127.0.0.2:2389
initial-cluster-token: etcd-cluster
initial-cluster:
etcd-1=http://127.0.0.2:2390,etcd-2=http://127.0.0.2:2391,etcd-3=h
ttp://127.0.0.2:2392
initial-cluster-state: new
```

In the same folder as the one you run the commands from, create three data directories, namely etcd1, etcd2 and etcd3. In my case, these will be created in the /tmp directory.

Open up a terminal window, and type the command

```
sudo etcd-download-test/etcd --config-file
<path_to_etcd1.conf.yml>
```

Similarly, open two more windows and type the same command with config files 2 and 3. Once all three commands are running, it should look like this -



```
{"level":"info","ts":"2024-03-21T17:24:01.76196+0530","caller":"rafthttp/stream.
go:249","msg":"set message encoder","from":"80f0dfea5ea0c542","to":"f42c8c852ecd
aa90","stream-type":"stream MsgApp v2"}
{"level":"info","ts":"2024-03-21T17:24:01.762227+0530","caller":"rafthttp/stream
.go:274","msg":"established TCP streaming connection with remote peer","stream-w
riter-type":"stream MsgApp v2","local-member-id":"80f0dfea5ea0c542","remote-peer
-id":"f42c8c852ecdaa90"}
{"level":"info","ts":"2024-03-21T17:24:01.762681+0530","caller":"rafthttp/stream
.go:412","msg":"established TCP streaming connection with remote peer","stream-r
eader-type":"stream MsgApp v2","local-member-id":"80f0dfea5ea0c542","remote-peer
-id":"f42c8c852ecdaa90"}
{"level":"info","ts":"2024-03-21T17:24:01.762964+0530","caller":"rafthttp/stream
.go:412","msg":"established TCP streaming connection with remote peer","stream-r
eader-type":"stream Message","local-member-id":"80f0dfea5ea0c542","remote-peer-i
d":"f42c8c852ecdaa90"}
{"level":"warn","ts":"2024-03-21T17:24:01.850722+0530","caller":"rafthttp/probin
g_status.go:68","msg":"prober detected unhealthy status","round-tripper-name":"R
OUND_TRIPPER_RAFT_MESSAGE","remote-peer-id":"f42c8c852ecdaa90","rtt":"0s","error
":"dial tcp 127.0.0.2:2392: connect: connection refused"}
{"level":"warn","ts":"2024-03-
g_status.go:68","msg":"prober
OUND_TRIPPER_SNAPSHOT","remote
ial tcp 127.0.0.2:2392: connec
```

```
.go:274","msg":"established TCP streaming connection with remote peer","stream-w
riter-type":"stream Message","local-member-id":"366952c5ece74e7c","remote-peer-i
d":"f42c8c852ecdaa90"}
{"level":"info","ts":"2024-03-21T17:24:01.76261+0530","caller":"rafthttp/stream.
go:412","msg":"established TCP streaming connection with remote peer","stream-re
ader-type":"stream Message","local-member-id":"366952c5ece74e7c","remote-peer-id
":"f42c8c852ecdaa90"}
{"level":"info","ts":"2024-03-21T17:24:01.762845+0530","caller":"rafthttp/stream
.go:412","msg":"established TCP streaming connection with remote peer","stream-r
eader-type":"stream MsgApp v2","local-member-id":"366952c5ece74e7c","remote-peer
-id":"f42c8c852ecdaa90"}
{"level":"info","ts":"2024-03-21T17:24:01.78055+0530","caller":"etcdserver/serve
r.go:783","msg":"initialized peer connections; fast-forwarding election ticks","
local-member-id":"366952c5ece74e7c","forward-ticks":8,"forward-duration":"800ms"
,"election-ticks":10,"election-timeout":"1s","active-remote-members":2}
{"level":"warn","ts":"2024-03-21T17:24:01.883943+0530","caller":"rafthttp/probin
g_status.go:68","msg":"prober detected unhealthy status","round-tripper-name":"R
OUND_TRIPPER_SNAPSHOT","remote-peer-id":"f42c8c852ecdaa90","rtt":"0s","error":"d
ial tcp 127.0.0.2:2392: connect: connection refused"}
                                                  ller":"rafthttp/probin
                                                  round-tripper-name":"R
                                                  a90","rtt":"0s","error
```

```
oder","from":"f42c8c852ecdaa90","to":"366952c5ece74e7c","stream-type":"stream MsgApp v2"}
{"level":"info","ts":"2024-03-21T17:24:01.762944+0530","caller":"rafthttp/stream.go:274","msg":"established TCP
 streaming connection with remote peer","stream-writer-type":"stream MsgApp v2","local-member-id":"f42c8c852ecd
aa90","remote-peer-id":"366952c5ece74e7c"}
{"level":"info","ts":"2024-03-21T17:24:01.804028+0530","caller":"etcdserver/server.go:783","msg":"initialized p
eer connections; fast-forwarding election ticks","local-member-id":"f42c8c852ecdaa90","forward-ticks":8,"forwar
d-duration":"800ms","election-ticks":10,"election-timeout":"1s","active-remote-members":2}
{"level":"info","ts":"2024-03-21T17:24:01.814891+0530","caller":"embed/serve.go:103","msg":"ready to serve clie
nt requests"}
{"level":"info","ts":"2024-03-21T17:24:01.814894+0530","caller":"etcdserver/server.go:2068","msg":"published lo
cal member to cluster through raft","local-member-id":"f42c8c852ecdaa90","local-member-attributes":"{Name:etcd-
3 ClientURLs:[http://127.0.0.2:2389]}","request-path":"/0/members/f42c8c852ecdaa90/attributes","cluster-id":"12
c5471a43cc6e9c","publish-timeout":"7s"}
{"level":"info","ts":"2024-03-21T17:24:01.814909+0530","caller":"embed/serve.go:103","msg":"ready to serve clie
nt requests"}
{"level":"info","ts":"2024-03-21T17:24:01.81512+0530","caller":"etcdmain/main.go:44","msg":"notifying init daem
on"}
{"level":"info","ts":"2024-03-21T17:24:01.815164+0530","caller":"etcdmain/main.go:50","msg":"successfully notif
ied init daemon"}
{"level":"info","ts":"2024-03-21T17:24:01.816521+0530","caller":"embed/serve.go:187","msg":"serving client traf
fic insecurely; this is strongly discouraged!","traffic":"grpc+http","address":"127.0.0.1:2389"}
{"level":"info","ts":"2024-03-21T17:24:01.816605+0530","caller":"embed/serve.go:187","msg":"serving client traf
fic insecurely; this is strongly discouraged!","traffic":"grpc+http","address":"127.0.0.2:2389"}
```

Activities
Terminal
Mar 21 17:24
athira@athira-Vostro-3500: /tmp
athira@athira-Vostro-3500: /tmp
athira@athira-Vostro-3500: /tmp

Screenshot 3(a)

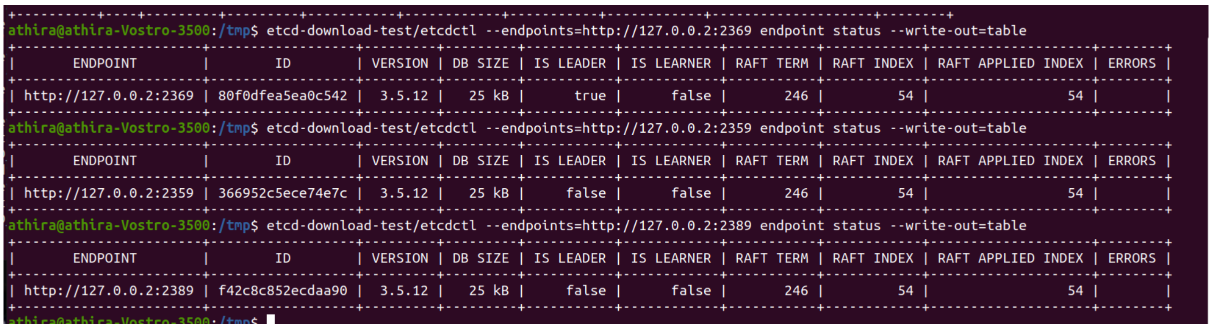
Note: If you get a permission denied error while starting the nodes, use sudo

E) Leader Election and Re-election

While the above three created nodes are running, in a fourth terminal, we will check the status of each end point and see which one is the leader.

Type in the commands -

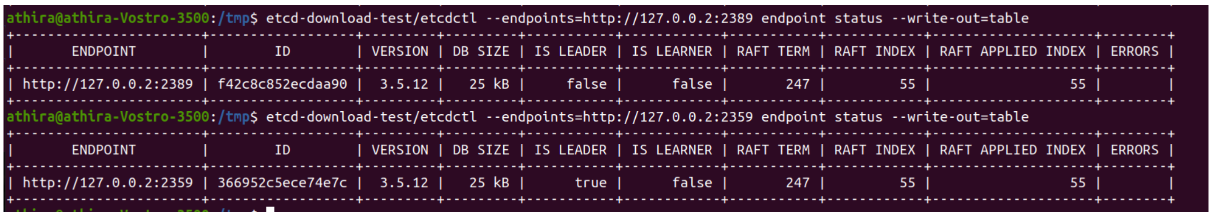
```
etcdctl --endpoints=http://127.0.0.2:2359 endpoint status
-write-out-table
etcdctl --endpoints=http://127.0.0.2:2369 endpoint status
-write-out-table
etcdctl --endpoints=http://127.0.0.2:2389 endpoint status
-write-out-table
```



```
athira@athira-Vostro-3500:/tmp$ etcd-download-test/etcdctl --endpoints=http://127.0.0.2:2369 endpoint status --write-out=table
+-----------------------+------------------+---------+---------+-----------+------------+-----------+------------+--------------------+--------+
|       ENDPOINT        |        ID        | VERSION | DB SIZE | IS LEADER | IS LEARNER | RAFT TERM | RAFT INDEX | RAFT APPLIED INDEX | ERRORS |
+-----------------------+------------------+---------+---------+-----------+------------+-----------+------------+--------------------+--------+
| http://127.0.0.2:2369 | 80f0dfea5ea0c542 |  3.5.12 |   25 kB |      true |      false |       246 |         54 |                 54 |        |
+-----------------------+------------------+---------+---------+-----------+------------+-----------+------------+--------------------+--------+
athira@athira-Vostro-3500:/tmp$ etcd-download-test/etcdctl --endpoints=http://127.0.0.2:2359 endpoint status --write-out=table
+-----------------------+------------------+---------+---------+-----------+------------+-----------+------------+--------------------+--------+
|       ENDPOINT        |        ID        | VERSION | DB SIZE | IS LEADER | IS LEARNER | RAFT TERM | RAFT INDEX | RAFT APPLIED INDEX | ERRORS |
+-----------------------+------------------+---------+---------+-----------+------------+-----------+------------+--------------------+--------+
| http://127.0.0.2:2359 | 366952c5ece74e7c |  3.5.12 |   25 kB |     false |      false |       246 |         54 |                 54 |        |
+-----------------------+------------------+---------+---------+-----------+------------+-----------+------------+--------------------+--------+
athira@athira-Vostro-3500:/tmp$ etcd-download-test/etcdctl --endpoints=http://127.0.0.2:2389 endpoint status --write-out=table
+-----------------------+------------------+---------+---------+-----------+------------+-----------+------------+--------------------+--------+
|       ENDPOINT        |        ID        | VERSION | DB SIZE | IS LEADER | IS LEARNER | RAFT TERM | RAFT INDEX | RAFT APPLIED INDEX | ERRORS |
+-----------------------+------------------+---------+---------+-----------+------------+-----------+------------+--------------------+--------+
| http://127.0.0.2:2389 | f42c8c852ecdaa90 |  3.5.12 |   25 kB |     false |      false |       246 |         54 |                 54 |        |
+-----------------------+------------------+---------+---------+-----------+------------+-----------+------------+--------------------+--------+
athira@athira-Vostro-3500:/tmp$
```

Screenshot 4(a)

In the above case, the leader is the node with port number 2369, as seen in the IS LEADER column of the table. Now, we will terminate the leader and see how re-election happens, i.e. the Raft consensus algorithm in action.
Go to the terminal where the leader node is running and terminate it (Ctrl+C). Wait for a few seconds, and obtain the endpoint statuses of all running nodes again (in this case, 2389 and 2359).



```
athira@athira-Vostro-3500:/tmp$ etcd-download-test/etcdctl --endpoints=http://127.0.0.2:2389 endpoint status --write-out=table
+-----------------------+------------------+---------+---------+-----------+------------+-----------+------------+--------------------+--------+
|       ENDPOINT        |        ID        | VERSION | DB SIZE | IS LEADER | IS LEARNER | RAFT TERM | RAFT INDEX | RAFT APPLIED INDEX | ERRORS |
+-----------------------+------------------+---------+---------+-----------+------------+-----------+------------+--------------------+--------+
| http://127.0.0.2:2389 | f42c8c852ecdaa90 |  3.5.12 |   25 kB |     false |      false |       247 |         55 |                 55 |        |
+-----------------------+------------------+---------+---------+-----------+------------+-----------+------------+--------------------+--------+
athira@athira-Vostro-3500:/tmp$ etcd-download-test/etcdctl --endpoints=http://127.0.0.2:2359 endpoint status --write-out=table
+-----------------------+------------------+---------+---------+-----------+------------+-----------+------------+--------------------+--------+
|       ENDPOINT        |        ID        | VERSION | DB SIZE | IS LEADER | IS LEARNER | RAFT TERM | RAFT INDEX | RAFT APPLIED INDEX | ERRORS |
+-----------------------+------------------+---------+---------+-----------+------------+-----------+------------+--------------------+--------+
| http://127.0.0.2:2359 | 366952c5ece74e7c |  3.5.12 |   25 kB |      true |      false |       247 |         55 |                 55 |        |
+-----------------------+------------------+---------+---------+-----------+------------+-----------+------------+--------------------+--------+
```

Screenshot 4(b)

We can see that the current leader is node with port number 2359.

This brings us to the end of the assignment. Submit the above mentioned screenshots in a PDF file.