Tutorial Apache Mesos

Para realizar esse tutorial, é necessária uma máquina rodando Ubuntu server, podendo estar rodando na plataforma de sua escolha: AWS, Azure ou VirtualBox na sua própria máquina.

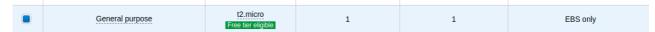
Nesse começo trataremos sobre como criar uma instância na AWS adequada para rodar o Mesos.

Criando instância na AWS EC2

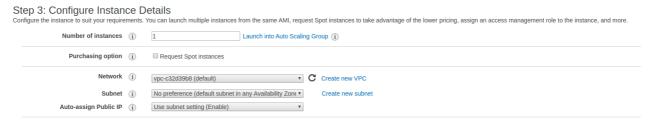
- 1. Realize o login no AWS Console e navegue para a seção do EC2.
- 2. Clique em Launch Instance e selecione a imagem Ubuntu Server 16.04 LTS AMI (Eligible for Free tier).



3. Em Instance Type selecione t2.micro e continue para Configure Instance Details.



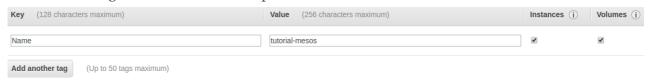
4. Selecione para todos os campos a opção **default** e continue para **Add Storage**.



5. Em storage mantenha a opção padrão: 8GB General Purpose SSD.



6. Adicione uma tag de **Name** na instância para indentifica-lá.



7. Na seção de **Security Groups**, crie um com as portas **22**, **2181**, **5050** abertas. Caso planeje instalar o Marathon posteriormente, abra a porta 8080 também para acessar o Marathon UI.

Type (j)	Protocol (j)	Port Range (j)	Source (j)
Custom TCP F ▼	TCP	8080	Custom • 0.0.0.0/0, ::/0
Custom TCP F ▼	TCP	5050	Custom v 0.0.0.0/0
SSH ▼	TCP	22	Custom • 0.0.0.0/0
Custom TCP F ▼	TCP	2181	Custom v 0.0.0.0/0

- 8. Por fim, clique em **Review and Launch** e selecione uma key pair (caso já possua uma) ou crie e baixe uma para acessar a instância.
- 9. Assim que a instância estiver rodando, acesse-a por meio de uma conexão SSH utilizando o Public IP e a key pair existente ou criada.

```
Public DNS (IPv4) ec2-34-201-73-112.compute-1.amazonaws.com

IPv4 Public IP 34.201.73.112

IPv6 IPs -
```

- 10. Se você estiver usando Windows, será necessário utilizar o Putty para gerar um arquivo .ppk a partir do arquivo baixado .pem e para acessar a instância.
- 11. Se estiver usando Linux, é possível se conectar diretamente do terminal utilizando o arquivo .pem e o Public IP da instância. Para isso basta utilizar os seguintes comandos, subtituindo o [/dir/da/keypair.pem] pelo diretório onde o arquivo .pem da keypair está :

```
$ chmod 400 /dir/da/keypair.pem
$ ssh -i /dir/da/keypair.pem ubuntu@[IP-DA-INSTANCIA]
```

Configurando ambiente da Instância

1. Após se conectar a instância, é preciso instalar o Java JDK 8.

```
$ sudo apt-get update
$ sudo apt install openjdk-8-jdk openjdk-8-jre
```

Então, execute os seguintes comandos para adicionar as chaves OpenPGP para o pacote do Mesos.
 Caso o primeiro comando falhe, tente com o segundo.

```
$ sudo apt-key adv --keyserver keyserver.ubuntu.com --recv DF7D54CBE56151BF
OU
$ sudo apt-key adv --keyserver keyserver.ubuntu.com --recv E56151F
```

3. Depois adicione o repositório do Mesos especifico para a versão do Ubuntu sendo utilizada, como estamos utilizando a versão 16.04, seu codinome é **Xenial**, como escrito abaixo.

```
$ sudo sh -c 'echo "deb http://repos.mesosphere.com/ubuntu xenial main" >> /et
c/apt/sources.list.d/mesosphere.list'
$ sudo apt-get update
```

4. Caso tudo tenha ocorrido bem, podemos agora instalar o Mesos do pacote adicionado.

```
$ sudo apt-get -y install mesos
```

Antes de começar a rodar o Mesos é preciso configurar algumas variáveis e arquivos.

Configurando o Mesos

1. Execute o comando:

```
$ hostname -f
```

Para obter o hostname da Instância onde o Mesos Master será iniciado.

```
ubuntu@ip-172-31-90-68:/etc/zookeeper/conf$
ubuntu@ip-172-31-90-68:/etc/zookeeper/conf$
ubuntu@ip-172-31-90-68:/etc/zookeeper/conf$
ubuntu@ip-172-31-90-68:/etc/zookeeper/conf$
ubuntu@ip-172-31-90-68:/etc/zookeeper/conf$ hostname -f
ip-172-31-90-68.ec2.internal
ubuntu@ip-172-31-90-68:/etc/zookeeper/conf$
```

2. Navegue para o diretório: /etc/zookeeper/conf e edite o arquivo zoo.cfg como mostrado abaixo. Descomente a linha que se inicia com server.1 e substitua zookeeper1 com o hostname obtido anteriormente. Como estamos utilizando um zookeeper em um única máquina, podemos substituir por localhost.

```
$ cd /etc/zookeeper/conf
$ sudo nano zoo.cfg
```

```
GNU nano 2.5.3
                                File: zoo.cfg
# the directory where the snapshot is stored.
dataDir=/var/lib/zookeeper
# Place the dataLogDir to a separate physical disc for better performance
# dataLogDir=/disk2/zookeeper
# the port at which the clients will connect
clientPort=2181
# specify all zookeeper servers
 The fist port is used by followers to connect to the leader
# The second one is used for leader election
server.1=ip-172-31-90-68.ec2.internal:2888:3888
#server.2=zookeeper2:2888:3888
#server.3=zookeeper3:2888:3888
# To avoid seeks ZooKeeper allocates space in the transaction log file in
# blocks of preAllocSize kilobytes. The default block size is 64M. One reason
 for changing the size of the blocks is to reduce the block size if snapshots
                                [ Wrote 52 lines ]
             ^O Write Out ^W Where Is
                                        ^K Cut Text
                                                     ^J Justify
^G Get Help
                                                                   ^C Cur Pos
   Exit
                Read File ^\
                             Replace
                                          Uncut Text<sup>^</sup>T
                                                                     Go To Line
                                                        To Spell
```

3. Rode o seguinte comando para editar o arquivo **myid** dentro do diretório /**etc/zookeeper/conf**. Aqui estamos adicionando o número 1 ao arquivo. Esse número é o server id para cada instância zookeeper do cluster, ele não pode ser repetido e varia de 1 a 255.

```
$ cd /etc/zookeeper/conf
$ sudo sh -c 'echo -n "1" >> myid'
```

4. Navegue para o diretório: /etc/mesos e edite o arquivo zk para apontar para o IP interno da instância do zookeeper como mostrado na imagem abaixo, novamente, como estamos utilizando uma única máquina pode-se usar localhost. Caso haja mais de uma instância rodando um mesos-master, o mesmo deverá ser feito em todas.

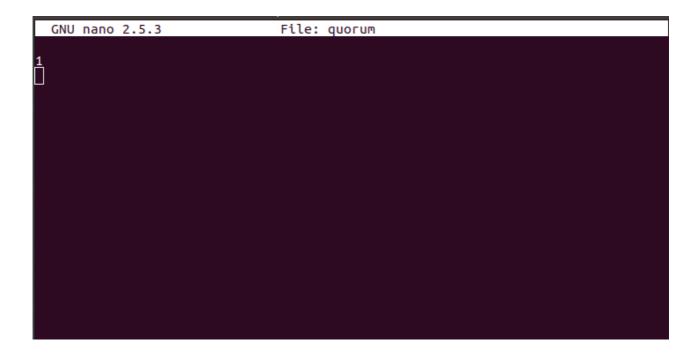
```
$ cd /etc/mesos
$ sudo nano zk
```

```
GNU nano 2.5.3 File: zk

zk://localhost:2181/mesos
```

5. Em seguida, navegue para o diretório: /etc/mesos-master e edite o aquivo quorum com um valor. Esse valor deve ser definido de uma maneira que 50% do masters devem estar "presentes" para que uma decisão seja tomada. Assim, ele deve ser maior que o número de masters divido por 2, como estamos rodando apenas uma instância com um master, definimos como 1.

```
$ cd /etc/mesos-master
$ sudo nano quorum
```



6. Ainda é preciso configurar o hostname e endereco IP para nosso mesos-master. Crie 2 arquivos **ip** e **hostname**, dentro do diretorio /**etc/mesos-master** com os valores corretos.

```
$ cd /etc/mesos-master
$ sudo sh -c 'echo "[IP-INTERNO-DA-INSTANCIA]" >> ip'
$ sudo sh -c 'echo "[IP-INTERNO-DA-INSTANCIA]" >> hostname'
$ cat ip
$ cat hostname
```

```
ubuntu@ip-172-31-90-68:/etc/mesos-master$
ubuntu@ip-172-31-90-68:/etc/mesos-master$
ubuntu@ip-172-31-90-68:/etc/mesos-master$
ubuntu@ip-172-31-90-68:/etc/mesos-master$ cat ip
172.31.90.68
ubuntu@ip-172-31-90-68:/etc/mesos-master$ cat hostname
172.31.90.68
ubuntu@ip-172-31-90-68:/etc/mesos-master$
```

Observação: Caso o mesos seja executado em uma subnet privada, utilize o IP Interno da instância para ambos os arquivos. Mas caso esteja em uma subnet pública, utilize o IP público da instância.

Iniciando o zookeeper, mesos-master e mesos-slave

Agora que o ambiente e as variáveis estão todas configuradas, podemos iniciar o processo do mesos-master, mesos-slave e do zookeeper. Para isso, utilize os seguintes comandos:

```
$ sudo service zookeeper start
$ sudo service mesos-master start
$ sudo service mesos-slave start
```

Para verificar que tudo está rodando sem erros utilize os seguintes comandos para obter o status dos processos:

```
$ sudo service zookeeper status
$ sudo service mesos-master status
$ sudo service mesos-slave status
```

O output deve ser parecido com o mostrado abaixo para o zookeeper:

Para o mesos-master:

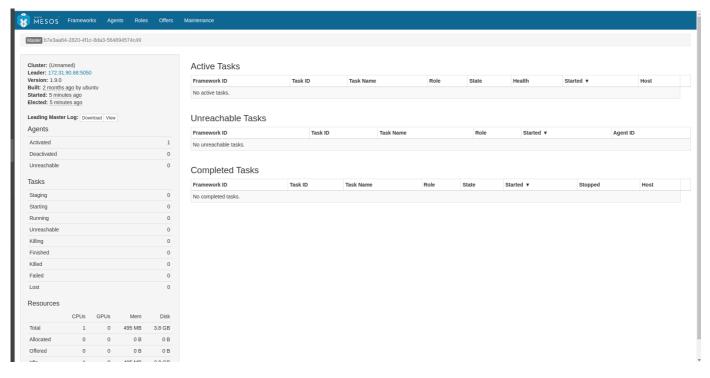
```
| Presos-master service - Mesos Master | Loaded: Loaded (/lib/system/nesos-master.service; disabled; vendor preset: enabled) | Active: active (running) since Qua 2019-11-13 14:34:01 UTC; Imin 23s ago | Main PID: 1843 (neoss-master) | Main PID: 1843 (neoss-master) | Main PID: 18443 (neoss-master)
```

Para o mesos-slave:

Quando todos os serviços estiverem rodando corretamente, utilize seu browser e acesse o endereço [IP_PUBLICO_INSTANCIA]:5050 para entrar no dashboard do mesos, ilustrado abaixo:

Observação: Se seu computador estiver conectado a uma rede pública, como de faculdades, esse endereço

pode ser bloqueado no browser.



Como não rodamos nenhum framework ainda, a lista de frameworks estará vazia.

Para executar um comando simples na infraestrutura recém-criada, para testá-la podemos utilizar a seguinte sintaxe:

```
$ mesos-execute --master=<IP_PUBLICO>:5050 --name="echo-test" --command=echo "Hello, W
orld"
```

```
ubuntugitp-172-31-90-68:-$ mesos-execute --master=34.201.73.112:5050 --name="echo-test" --command=echo "Hello, World"
11113 15:01:34.161592 1761 scheduler.cpp:342] Using default 'basic' HTTP authenticatee
1113 15:01:34.165376 1769 scheduler.cpp:342] Using default 'basic' HTTP authenticatee
1113 15:01:34.165376 1769 scheduler.cpp:342] New master detected at master@34.201.73.112:5050
Subscribed with 10 Dresanad-4:280-4616-sdas-5-66984547c49-0808 5ubnitted task 'echo-test' to agent '94987e3b-fae-4c80-ab5d-30473e071c53-50'
Received status update TASK_STARTING for task 'echo-test' source: SOURCE_EXECUTOR
Received status update TASK_STARTING for task 'echo-test' source: SOURCE_EXECUTOR
Received status update TASK_STINISHED for task 'echo-test' nessage: 'Command extited with status 0'
source: SOURCE_EXECUTOR
Received status update TASK_STINISHED for task 'echo-test' source: SOURCE_EXECUTOR
ubuntugitp-172-31-90-68:-$ 

Ubuntugitp-172-31-90-68:-$ 

Ubuntugitp-172-31-90-68:-$
```

Criando um framework

Como explicado anteriormente, para executar tarefas sobre a infraestrutura gerenciada pelo mesos master é preciso de um framework que define essa intermediação. Assim, vamos explicar abaixo a composição e configuração de um framework em **Python**.

Para esse framework utilizamos a biblioteca **pymesos** disponível para Python e realizamos os seguintes imports no arquivo python:

```
from pymesos import MesosSchedulerDriver, Scheduler, encode_data
import uuid
from addict import Dict
import socket
import getpass
from threading import Thread
import signal
import time
import enum
import os
```

Primeiramente temos uma classe que define as **Tasks**, estas recebem em sua inicialização um *taskId*, um *comando* a ser executado e os **resources necessários (cpu e memória)**.

```
class Task:
    def __init__(self, taskId, command, cpu, mem):
        self.taskId = taskId
        self.command = command
        self.cpu = cpu
        self.mem = mem
```

Então temos um conjunto de métodos para essa classe que são utilizados para aceitar uma oferta de recursos para executar a tarefa:

• O método **__getResource** retorna valor de um recurso específico disponibilizado em uma oferta, por exemplo, **memória**, ele é utilizado na etapa de definir se os recursos de uma oferta são suficientes para rodar a tarefa.

```
def __getResource(self, res, name):
    for r in res:
        if r.name == name:
            return r.scalar.value
    return 0.0
```

• O método __updateResource atualiza o valor de um recurso específico, ele é utilizado quando uma oferta é aceita. Assim, ele basicamente subtrai o valor necessário de certo recurso para a tarefa do valor total de certo recurso na oferta.

```
def __updateResource(self, res, name, value):
    if value <= 0:
        return
    for r in res:
        if r.name == name:
            r.scalar.value -= value
    return</pre>
```

• O método **acceptOrder** determina se uma oferta recebida é adequada para executar a tarefa. Assim, a partir dos métodos acima ele verifica se o valor de recursos da oferta é maior que o valor de recursos necessários para a execução da tarefa e aceita ou não esta oferta, atualizando os recursos disponíveis caso a oferta seja aceitada.

```
def acceptOffer(self, offer):
   accept = True
   if self.cpu != 0:
        cpu = self.__getResource(offer.resources, "cpus")
        if self.cpu > cpu:
            accept = False
    if self.mem != 0:
        mem = self.__getResource(offer.resources, "mem")
        if self.mem > mem:
            accept = False
    if(accept == True):
        self.__updateResource(offer.resources, "cpus", self.cpu)
        self.__updateResource(offer.resources, "mem", self.mem)
        return True
    else:
        return False
```

Em seguida temos uma classe que define o **Scheduler**, ela é inicializada com listas e dicionários vazios que são usados para armazenarem as tarefas disponíveis, em inicialização, em execução e terminadas.

Além disso, neste caso na incialização da classe já inserimos na lista de tarefas disponíveis as tarefas a serem executadas, mas poderiamos criar um método para que a classe recebesse as tarefas dinâmicamente. Essas tarefas serão melhor explicadas mais a frente neste guia.

Então, temos um método *resourceOffers* que verifica se há alguma tarefa pendente na lista e caso haja, ele cria uma oferta para essa tarefa por meio do método *acceptOffer* da classe **Task**. Assim, caso a oferta seja aceita ele executa a tarefa no cluster e a adiciona na lista correspondente.

```
def resourceOffers(self, driver, offers):
    logging.debug("Received new offers")
    logging.info("Recieved resource offers: {}".format([o.id.value for o in offer
s]))

if(len(self.idleTaskList) == 0):
    driver.suppressOffers()
    logging.info("Idle Tasks List Empty, Suppressing Offers")
    return

filters = {'refuse_seconds': 1}

for offer in offers:
    taskList = []
    pendingTaksList = []
    while True:
    if len(self.idleTaskList) == 0:
```

```
break
                Task = self.idleTaskList.pop(0)
                if Task.acceptOffer(offer):
                    task = Dict()
                    task_id = Task.taskId
                    task.task_id.value = task_id
                    task.agent_id.value = offer.agent_id.value
                    task.name = 'task {}'.format(task_id)
                    task.command.value = Task.command
                    task.resources = [
                        dict(name='cpus', type='SCALAR', scalar={'value': Task.cpu}),
                        dict(name='mem', type='SCALAR', scalar={'value': Task.mem}),
                    self.startingTaskList[task_id] = Task
                    taskList.append(task)
                    logging.info("Starting task: %s, in node: %s" % (Task.taskId, offe
r.hostname))
                else:
                    pendingTaksList.append(Task)
            if(len(taskList)):
                    driver.launchTasks(offer.id, taskList, filters)
            self.idleTaskList = pendingTaksList
```

Por fim, há o método **statusUpdate** que é utilizado para logar o status de execução da tarefa e atualizar sua lista e estado.

```
(update.task id.value, update.container status.network infos[0].ip add
resses[0].ip_address))
                self.runningTaskList[update.task_id.value] = Task
                del self.startingTaskList[update.task_id.value]
        elif update.state == "TASK_FAILED":
            Task = None
            if update.task_id.value in self.startingTaskList:
                Task = self.startingTaskList[update.task_id.value]
                del self.startingTaskList[update.task_id.value]
            elif update.task_id.value in self.runningTaskList:
                Task = self.runningTaskList[update.task_id.value]
                del self.runningTaskList[update.task_id.value]
            if Task:
                logging.info("Uni task: %s failed." % Task.taskId)
                self.idleTaskList.append(Task)
                driver.reviveOffers()
            else:
                logging.error("Received task failed for unknown task: %s" % update.tas
k_id.value )
        else:
            logging.info("Received status %s for task id: %s" % (update.state, update.
task_id.value))
```

O código completo do **scheduler.py** está presente aqui

Descrevendo agora as tarefas que foram criada na inicialização da classe do **Scheduler**, criamos duas tarefas a serem executadas sobre os agentes do mesos:

```
Task("taskHelloWorld", "echo HelloWorld", .1, 100)
Task("taskDIR", "mkdir /home/ubuntu/HelloMesos", .1, 100)
```

A primeira simplesmente executa um comando "*echo HelloWorld*", porém não é possível observar o output do comando, uma vez que ele roda no agente. Então, criamos uma segunda tarefa para conseguirmos observar de fato sua execução, nela há um comando para que uma pasta *HelloMesos* seja criada no diretório /*home/ubuntu*. É importante notar que na infraestrutura do Mesos que criamos nesse tutorial tanto o master

quanto os agentes estão na mesma máquina, por isso conseguimos observar o resultado do comando dessa segunda tarefa.

Finalmente, para rodar o framework criado basta criar um arquivo scheduler.py na máquina onde o tutorial de implantação foi executado ou clonar diretamente do github e configurar uma variável de ambiente com o IP público da máquina do Mesos Master:

```
$ nano scheduler.py
OU
$ git clone https://github.com/elijose55/Mesos-Tutorial.git
$ export MESOS_MASTER_IP=34.201.73.112
$ python3 scheduler.py
```

Então, após executar o **scheduler** você deverá receber um output como abaixo:

```
ubuntu@ip-172-31-90-68:~$ python3 scheduler.py
Scheduler running, Ctrl+C to quit.
DEBUG:root:Received new offers
INFO:root:Recieved resource offers: ['b7e3aa64-2820-4f1c-8da3-564894574c49-039']
INFO:root:Starting task: taskHelloWorld, in node: ip-172-31-90-68.ec2.internal
INFO:root:Starting task: taskDIR, in node: ip-172-31-90-68.ec2.internal
DEBUG:root:Task taskHelloWorld is starting.
DEBUG:root:Task taskDIR is starting.
INFO:root:Task taskHelloWorld running in 172.31.90.68. Moving to running list
INFO:root:Task taskDIR running in 172.31.90.68. Moving to running list
INFO:root:Received status TASK_FINISHED for task id: taskDIR
INFO:root:Received status TASK_FINISHED for task id: taskHelloWorld
DEBUG:root:Received new offers
INFO:root:Received resource offers: ['b7e3aa64-2820-4f1c-8da3-564894574c49-040']
INFO:root:Idle Tasks List Empty, Suppressing Offers
```

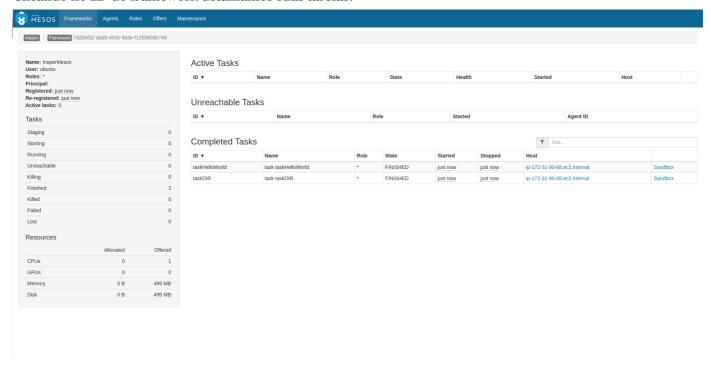
E como se pode observar o diretório *HelloMesos* é criado após a execução da tarefa no scheduler:

```
ubuntu@ip-172-31-90-68:~$ ls
HelloMesos scheduler.py
```

Além disso, é possível ver o framework criado na dashboard do Mesos, com seus detalhes e tarefas executadas.



Clicando no ID do framework detalhamos suas tarefas:



O framework criado foi baseado nas seguintes fontes:

http://mesos.apache.org/documentation/latest/app-framework-development-guide/

https://github.com/mesosphere/RENDLER/tree/master/python

https://github.com/smurli/pymesos-sample

O tutorial se baseou na seguinte fonte:

https://linuxacademy.com/guide/25034-introduction-to-apache-mesos/

Por Eli Jose, Pedro Azambuja e Rafael Viera