

Creation of key value pair

Successfully created key pair

Key pairs (1)

Filter key pairs

< 1 >

<input type="checkbox"/>	Name	Fingerprint	ID
<input type="checkbox"/>	second_pair	9d:27:0c:2c:07:ed:82:d0:f8:b4:45:15:0...	key-0b1bb4734bc1f171c

Cluster Creation

Clone Terminate AWS CLI export

Cluster: Assignment2Cluster_Sep09 Running Running step

Summary

Application user interfaces

Monitoring

Hardware

Configurations

Events

Steps

Bootstrap actions

Summary

ID: j-1UTN42PUVD8FE

Creation date: 2020-09-09 23:35 (UTC-5)

Elapsed time: 10 minutes

After last step completes: Cluster waits

Termination protection: Off [Change](#)

Tags: -- [View All](#) / [Edit](#)

Master public DNS: ec2-3-92-222-155.compute-1.amazonaws.com [Connect to the Master Node Using SSH](#)

Configuration details

Release label: emr-5.30.1

Hadoop distribution: Amazon 2.8.5

Applications: Hive 2.3.6, Hue 4.6.0, Mahout 0.13.0, Pig 0.17.0, Tez 0.9.2

Log URI: s3://aws-logs-975405783468-us-east-1/elasticmapreduce/

EMRFS consistent view: Disabled

Custom AMI ID: --

Application user interfaces

Persistent user interfaces: [YARN timeline server](#), [Tez UI](#)

On-cluster user interfaces: Not Enabled [Enable an SSH Connection](#)

Network and hardware

Availability zone: us-east-1a

Subnet ID: subnet-81822ba0 [View](#)

Master: Running 1 m4.large

Core: Running 1 m4.large

Task: --

Cluster scaling: Not enabled

Security and access

Key name: second_pair

EC2 instance profile: EMR_EC2_DefaultRole

EMR role: EMR_DefaultRole

Visible to all users: All [Change](#)

Security groups for Master: sg-0a95617f3bd132a6a [View](#) (ElasticMapReduce-master)

Security groups for Core & Task: sg-0b2fb7b2f354ac4da [View](#) (ElasticMapReduce-slave)

Selecting Master Security Group

Security Groups (1/2) Info

Filter security groups

search: sg-0a95617f3bd132a6a [Clear filters](#)

☒

☐

Name	Security group ID	Security group name	VPC ID	Description	Owner
-	sg-0a95617f3bd132a6a	ElasticMapReduce-master	vpc-424fb03f View	Master group for Elastic...	975...
- View	sg-0b2fb7b2f354ac4da	ElasticMapReduce-slave	vpc-424fb03f View	Slave group for Elastic ...	975...

Uploading the file to Hadoop (Akshay.txt)

```
hadoop@ip-172-31-94-70:~  
https://aws.amazon.com/amazon-linux-2/  
33 package(s) needed for security, out of 89 available  
Run "sudo yum update" to apply all updates.  
  
EEEEEEEEEEEEEEEEEEEE MMMMMMM      MMMMMMM RRRRRRRRRRRRRRR  
E::::::::::::::::::::E M:::::M      M:::::M R::::::::::::R  
EE::::::::::::::::::::E M:::::M      M:::::M R::::RRRRR::::R  
E:::E      EEEEE M:::::M      M:::::M RR:::R      R:::R  
E:::E      M:::::M      M:::M M:::::M R:::R      R:::R  
E::::EEEEEEEEEE M:::::M M:::M M:::M M:::::M R::RRRRR::::R  
E::::::::::::::::E M:::::M M:::M M:::::M R:::::::::RR  
E::::EEEEEEEEEE M:::::M M:::::M M:::::M R::::RRRRR::::R  
E:::E      M:::::M M:::M M:::::M R:::R      R:::R  
E:::E      EEEEE M:::::M      MMM      M:::::M R:::R      R:::R  
EE::::::::::::E M:::::M      M:::::M R:::R      R:::R  
E::::::::::::::::E M:::::M      M:::::M RR:::R      R:::R  
EEEEEEEEEEEEEEEEEEEE MMMMMMM      MMMMMMM RRRRRRR      RRRRRR  
  
[hadoop@ip-172-31-94-70 ~]$ ls  
[hadoop@ip-172-31-94-70 ~]$ ls  
akshay.txt  
[hadoop@ip-172-31-94-70 ~]$ |
```

S3 bucket creation

We've temporarily re-enabled the previous version of the S3 console while we continue to improve the new S3 console experience. [Switch to the new console.](#)

S3 buckets [Discover the console](#)

Search for buckets All access types

[+ Create bucket](#) [Edit public access settings](#) [Empty](#) [Delete](#)

2 Buckets 1 Regions

<input type="checkbox"/> Bucket name	Access	Region	Date created
<input checked="" type="checkbox"/> a20458999assignment2	Bucket and objects not public	US East (N. Virginia)	Sep 10, 2020 12:19:29 AM GMT-0500

Object creation in bucket

a20458999assignment2

Overview Properties Permissions Management Access points

Search: Type a prefix and press Enter to search. Press ESC to clear.

[Upload](#) [+ Create folder](#) [Download](#) [Actions](#)

US East (N. Virginia)

Name	Last modified	Size	Storage class
<input type="checkbox"/> A20458999.txt	Sep 10, 2020 12:26:12 AM GMT-0500	29.0 B	Standard

Viewing 1 to 1

09) Listing the files under root directory

Command Used :Hadoop fs -ls /

```
[hadoop@ip-172-31-94-70 ~]$ hadoop fs -ls /  
Found 4 items  
drwxr-xr-x - hdfs hadoop 0 2020-09-10 04:41 /apps  
drwxrwxrwt - hdfs hadoop 0 2020-09-10 04:42 /tmp  
drwxr-xr-x - hdfs hadoop 0 2020-09-10 04:41 /user  
drwxr-xr-x - hdfs hadoop 0 2020-09-10 04:41 /var  
[hadoop@ip-172-31-94-70 ~]$
```

10) List of files under user

Command Used: Hadoop fs -ls /user

```
[hadoop@ip-172-31-94-70 ~]$ hadoop fs -ls /user  
Found 6 items  
drwxrwxrwx - hadoop hadoop 0 2020-09-10 04:41 /user/hadoop  
drwxr-xr-x - mapred mapred 0 2020-09-10 04:41 /user/history  
drwxrwxrwx - hdfs hadoop 0 2020-09-10 04:41 /user/hive  
drwxrwxrwx - hue hue 0 2020-09-10 04:41 /user/hue  
drwxrwxrwx - oozie oozie 0 2020-09-10 04:41 /user/oozie  
drwxrwxrwx - root hadoop 0 2020-09-10 04:41 /user/root  
[hadoop@ip-172-31-94-70 ~]$
```

11)Creating new directory

Command Used: Hadoop fs -mkdir /user/csp554

```
drwxrwxrwx - root hadoop 0 2020-09-10 04:41 /user/root  
[hadoop@ip-172-31-94-70 ~]$ hadoop fs -mkdir /user/csp554  
[hadoop@ip-172-31-94-70 ~]$ hadoop fs -ls /user  
Found 7 items  
drwxr-xr-x - hadoop hadoop 0 2020-09-10 05:51 /user/csp554  
drwxrwxrwx - hadoop hadoop 0 2020-09-10 04:41 /user/hadoop  
drwxr-xr-x - mapred mapred 0 2020-09-10 04:41 /user/history  
drwxrwxrwx - hdfs hadoop 0 2020-09-10 04:41 /user/hive  
drwxrwxrwx - hue hue 0 2020-09-10 04:41 /user/hue  
drwxrwxrwx - oozie oozie 0 2020-09-10 04:41 /user/oozie  
drwxrwxrwx - root hadoop 0 2020-09-10 04:41 /user/root  
[hadoop@ip-172-31-94-70 ~]$
```

12)Creating second directory

Command Used: Hadoop fs -mkdir /user/csp554-2

```
[hadoop@ip-172-31-94-70 ~]$ hadoop fs -ls /user/csp554  
[hadoop@ip-172-31-94-70 ~]$ hadoop fs -mkdir /user/csp554-2  
[hadoop@ip-172-31-94-70 ~]$
```

13) Copy from local to user/csp554

Command Used: `hadoop fs -copyFromLocal /home/Hadoop/Akshay.txt /user/csp554`

```
[hadoop@ip-172-31-94-70 ~]$ ls
akshay.txt
[hadoop@ip-172-31-94-70 ~]$ hadoop fs -copyFromLocal /home/hadoop/akshay.txt /user/csp554
[hadoop@ip-172-31-94-70 ~]$ |
```

14) Copy from one directory to other

Command used: `Hadoop fs -cp /user/csp554/Akshay.txt /user/csp554-2`

```
-rw-r--r-- 1 hadoop hadoop 32 2020-09-10 06:22 /user/csp554/akshay.txt
[hadoop@ip-172-31-94-70 ~]$ hadoop fs -cp /user/csp554/akshay.txt /user/csp554-2
[hadoop@ip-172-31-94-70 ~]$ hadoop fs -ls /user/csp554-2
Found 1 items
-rw-r--r-- 1 hadoop hadoop 32 2020-09-10 06:30 /user/csp554-2/akshay.txt
[hadoop@ip-172-31-94-70 ~]$ |
```

15) Object copied from S3 bucket to hadoop(Renamed file as myid.txt).

Command Used: `aws s3 cp s3://a20458999assignment2/A20458999.txt /home/Hadoop/myid.txt`

Command used to list the newly copied file: `ls`

```
[hadoop@ip-172-31-94-70 ~]$ aws s3 cp s3://a20458999assignment2/A20458999.txt /home/hadoop/myid.txt
download: s3://a20458999assignment2/A20458999.txt to ./myid.txt
[hadoop@ip-172-31-94-70 ~]$ ls
akshay.txt myid.txt
[hadoop@ip-172-31-94-70 ~]$ |

[hadoop@ip-172-31-94-70 ~]$ aws s3 cp s3://a20458999assignment2/A20458999.txt /home/hadoop/a20458999.txt
download: s3://a20458999assignment2/A20458999.txt to ./a20458999.txt
[hadoop@ip-172-31-94-70 ~]$ ls
a20458999.txt akshay.txt myid.txt
[hadoop@ip-172-31-94-70 ~]$ |
```

16) Copy from s3 bucket to HDFS user/csp554-2 directory

Command Used: `Hadoop fs -cp s3://a20458999assignment2/A20458999.txt hdfs:///user/csp554-2`

```
[hadoop@ip-172-31-94-70 ~]$ hadoop fs -cp s3://a20458999assignment2/A20458999.txt hdfs:///user/csp554-2
20/09/10 06:51:28 INFO s3n.S3NativeFileSystem: Opening 's3://a20458999assignment2/A20458999.txt' for reading
[hadoop@ip-172-31-94-70 ~]$ hadoop fs -ls /user/csp554-2
Found 2 items
-rw-r--r-- 1 hadoop hadoop 29 2020-09-10 06:51 /user/csp554-2/A20458999.txt
-rw-r--r-- 1 hadoop hadoop 32 2020-09-10 06:30 /user/csp554-2/akshay.txt
[hadoop@ip-172-31-94-70 ~]$ |
```

17) Reading the content of the file.

Command Used: `hadoop fs -cat /user/csp554-2/A20458999.txt`

```
cat: /user/csp554-2/A20458999: No such file or directory
[hadoop@ip-172-31-94-70 ~]$ hadoop fs -cat /user/csp554-2/A20458999.txt
A20458999 this is the id file[hadoop@ip-172-31-94-70 ~]$ |
```

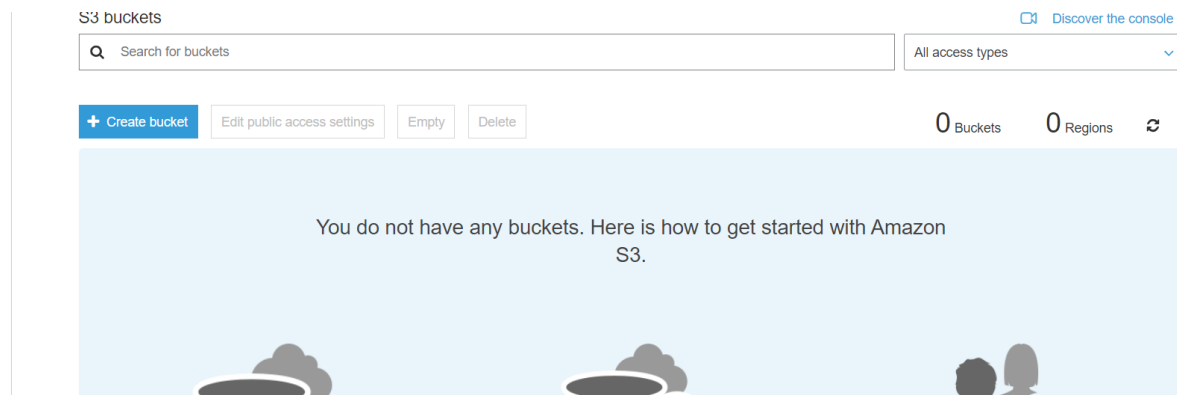
18) Deleting the file

Command Used: `hadoop fs -rm /user/csp554-2/A20458999.txt`

```
[hadoop@ip-172-31-94-70 ~]$ hadoop fs -rm /user/csp554-2/A20458999.txt
A20458999 this is the id file[hadoop@ip-172-31-94-70 ~]$ hadoop fs -rm /user/csp554-2/A20458999.txt
Deleted /user/csp554-2/A20458999.txt
[hadoop@ip-172-31-94-70 ~]$
```

```
[hadoop@ip-172-31-94-70 ~]$ hadoop fs -ls /user/csp554-2
Found 1 items
-rw-r--r-- 1 hadoop hadoop 32 2020-09-10 06:30 /user/csp554-2/akshay.txt
[hadoop@ip-172-31-94-70 ~]$
```

19) Cluster termination Bucket Deletion



Save up to 90% when running your EMR clusters with EC2 Spot Instances. [View tutorial](#)

[Create cluster](#) [View details](#) [Clone](#) [Terminate](#)

Filter: All clusters 2 clusters (all loaded)

	Name	ID	Status	Creation time (UTC-5)	Elapsed time	Normalized instance hours
<input type="checkbox"/>	Assignment2Cluster_Sep09	J-1UTN42PUVD8FE	Terminated User request	2020-09-09 23:35 (UTC-5)	2 hours, 29 minutes	24