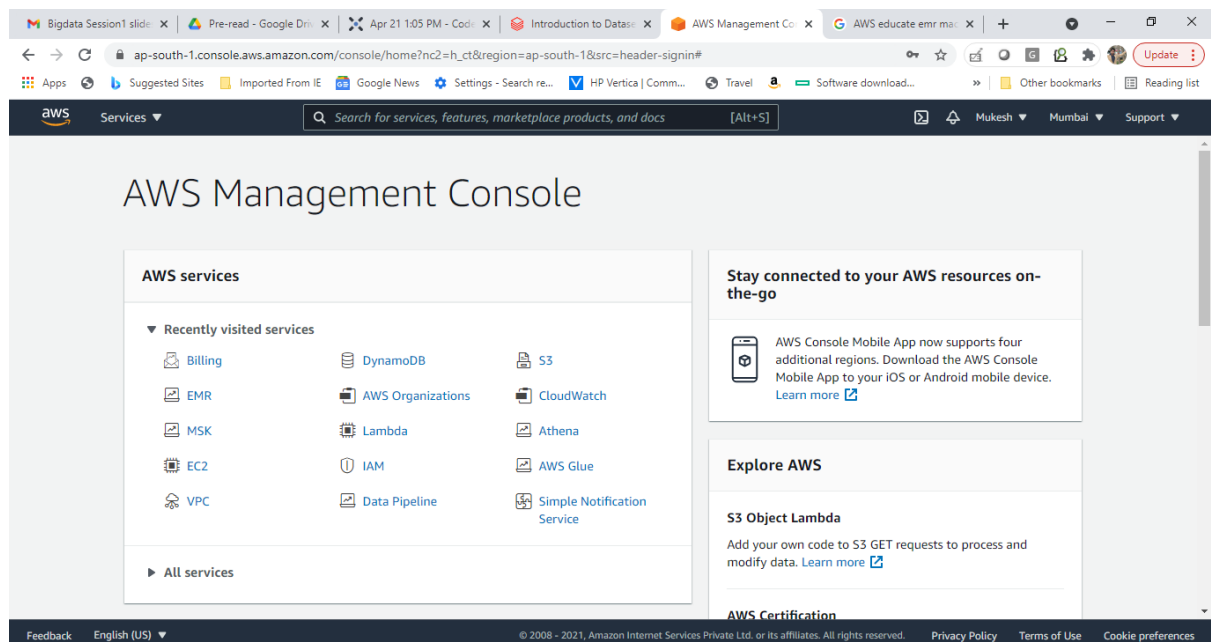
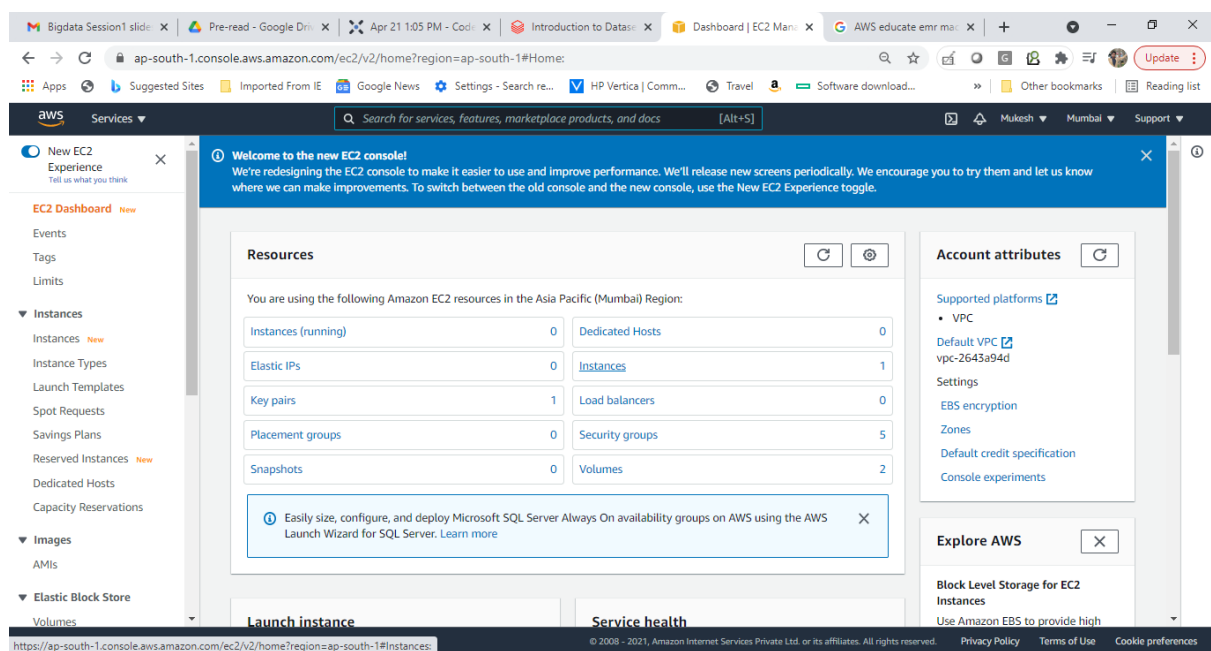


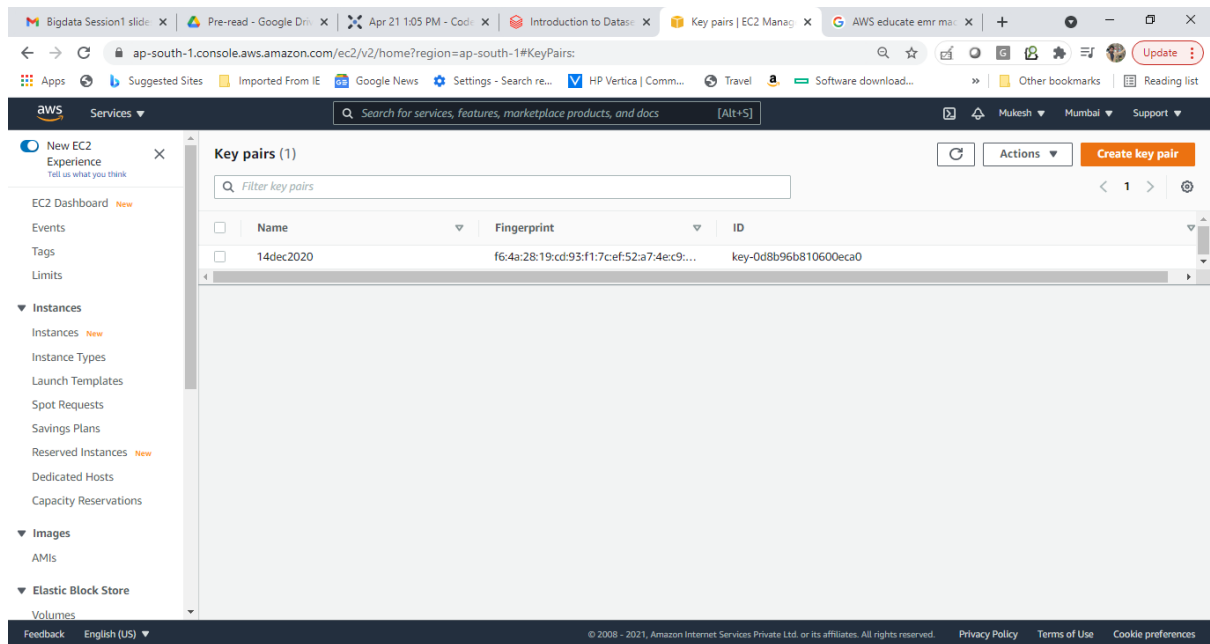
1. Login to your AWS account



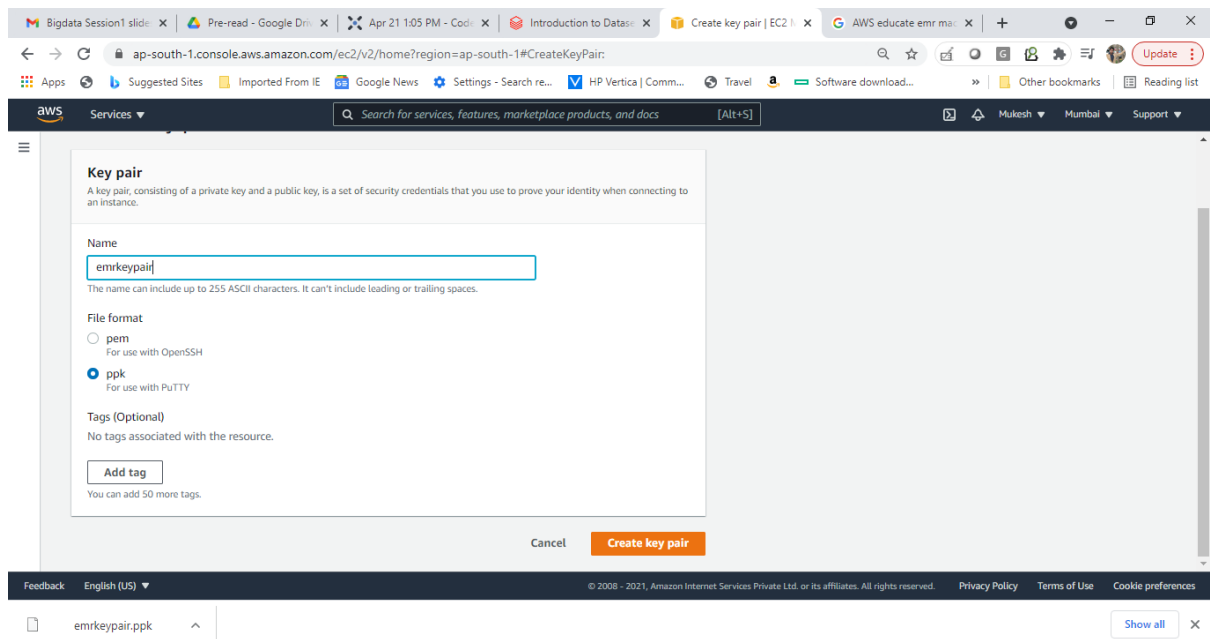
Here click on the “EC2” option which display page as shown below



Click on the “Key Pair” button on this screen which display below page.



On this page, click on the “Create key pair” option, it will take you to “Create key pair” page.



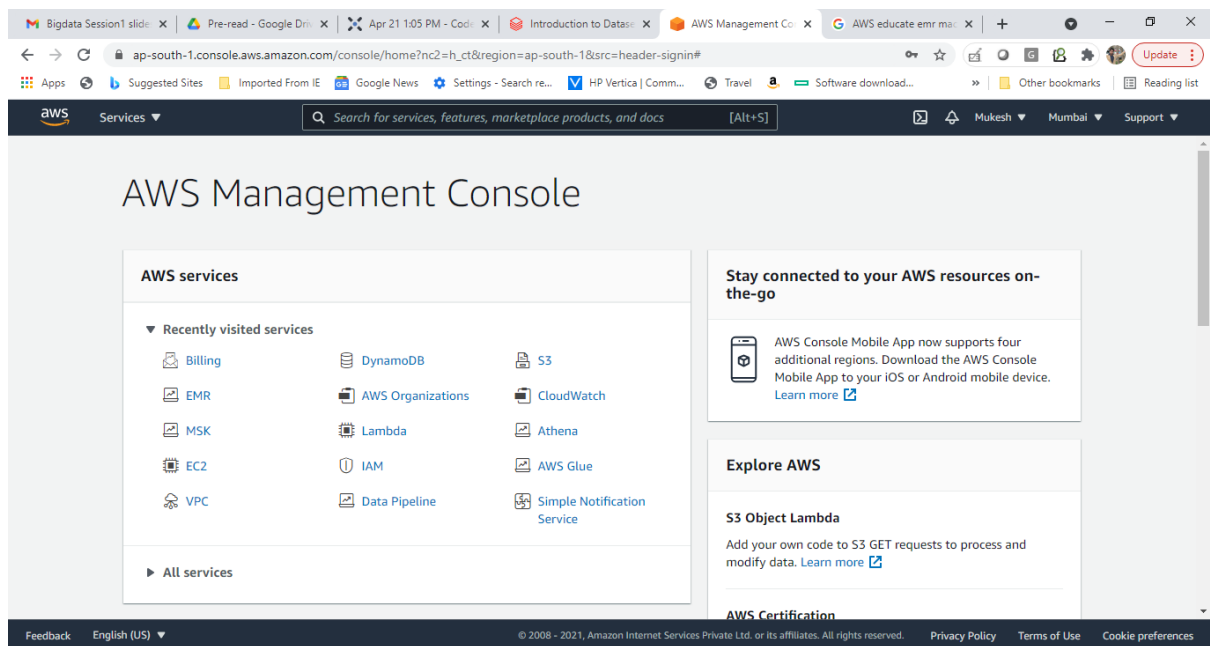
On the keypair is created it will download the ppk file and will display the page

This file is meant for personal use by rg.ravigupta91@gmail.com only.
Sharing or publishing the contents in part or full is liable for legal action.

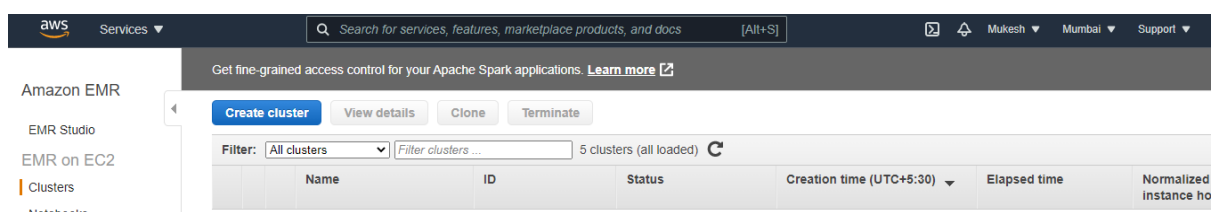
Key pairs (2)			
<input type="text" value="Filter key pairs"/>			
<input type="checkbox"/>	Name	Fingerprint	ID
<input type="checkbox"/>	14dec2020	f6:4a:28:19:cd:93:f1:7c:ef:52:a7:4e:c9:...	key-0d8b96b810600eca0
<input type="checkbox"/>	emrkeypair	7c:a8:f3:25:ab:73:ea:ff:7f:15:c6:c4:7c:0...	key-0058a2c105880eec0

2.

Click on the “aws” to go to the home page. You will see the below screen.



2. Click on the EMR , which will take you to “Create EMR Cluster” page.



3. Click on the “create cluster ” button which will take you to next page as shown below

The screenshot shows the AWS EMR console's 'Create Cluster - Quick Options' page. The 'General Configuration' section includes a 'Cluster name' field with 'My cluster' entered, a checked 'Logging' checkbox, and an 'S3 folder' field with 's3://aws-logs-953645729708-ap-south-1/elasticmapre'. The 'Launch mode' is set to 'Cluster'. The 'Software configuration' section shows the 'Release' label as 'emr-5.33.0'. Under 'Applications', 'Core Hadoop' is selected, listing Hadoop 2.10.1, Hive 2.3.7, Hue 4.9.0, Mahout 0.13.0, Pig 0.17.0, and Tez 0.9.2. Other application options like HBase, Presto, Spark, and Zeppelin are listed but not selected. A checkbox for 'Use AWS Glue Data Catalog for table metadata' is also present.

Here click on the “Go to Advance Options” button and select the options “radio buttons” shown in the below screen.

This screenshot shows the 'Create Cluster' page with advanced options. The 'Release' label is 'emr-5.33.0'. A grid of application checkboxes is displayed, with 'Hadoop 2.10.1', 'Hive 2.3.7', 'Hue 4.9.0', and 'Spark 2.4.7' selected. Other applications like Zeppelin, Tez, HBase, Presto, MXNet, Livy, Flink, Pig, ZooKeeper, Sqoop, Phoenix, and HCatalog are listed but not selected. Below the applications, there are sections for 'Multiple master nodes (optional)', 'AWS Glue Data Catalog settings (optional)' (with 'Use for Hive table metadata' and 'Use for Spark table metadata' checked), and 'Edit software settings' (with 'Enter configuration' selected). A text area for configuration is shown with the placeholder 'cClassification=config-file-name,properties=[myKey1=myValue1,myKey2=myValue2]'. The left sidebar shows the navigation menu with 'Step 3: General Cluster Settings' highlighted.

4.

Now select below “Concurrency” check box and “maximum steps that can run at a time” as 10 before clicking on the “Next” button

Steps (optional)

A step is a unit of work you submit to the cluster. For instance, a step might contain one or more Hadoop or Spark jobs. You can also submit additional steps to a cluster after it is running. [Learn more](#)

Concurrency: ☒ Run multiple steps at the same time to improve cluster utilization

maximum steps that can run at a time

Choose between 2 and 256 steps

After last step completes: ☒ Clusters enters waiting state

☐ Cluster auto-terminates

Step type

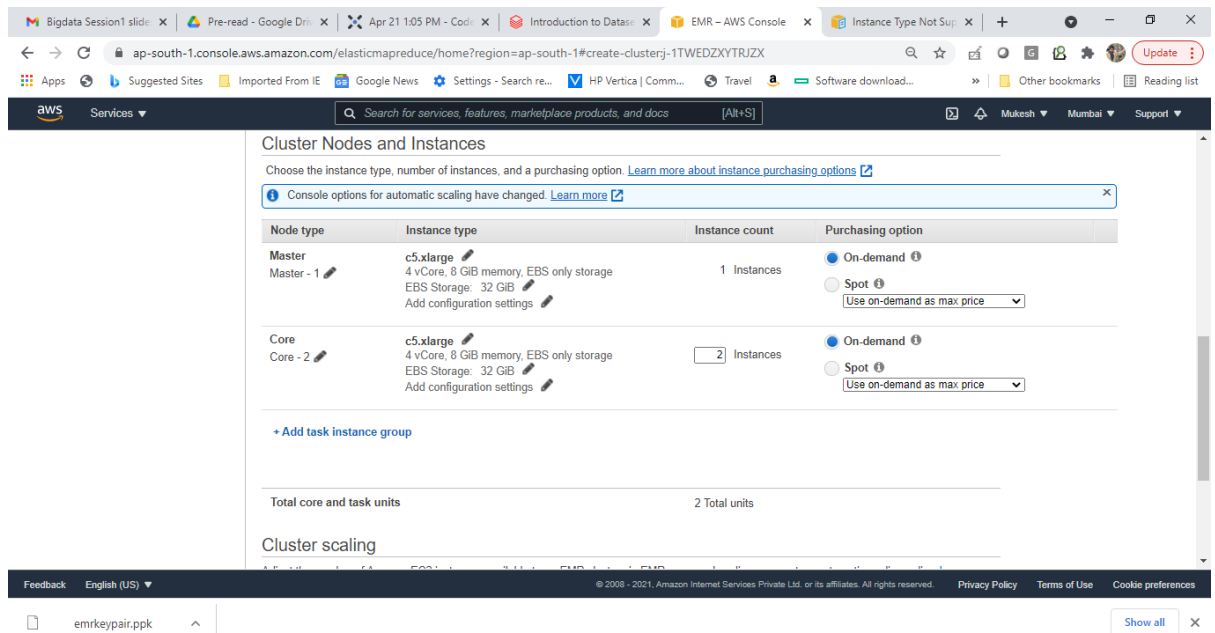
Add step

Cancel

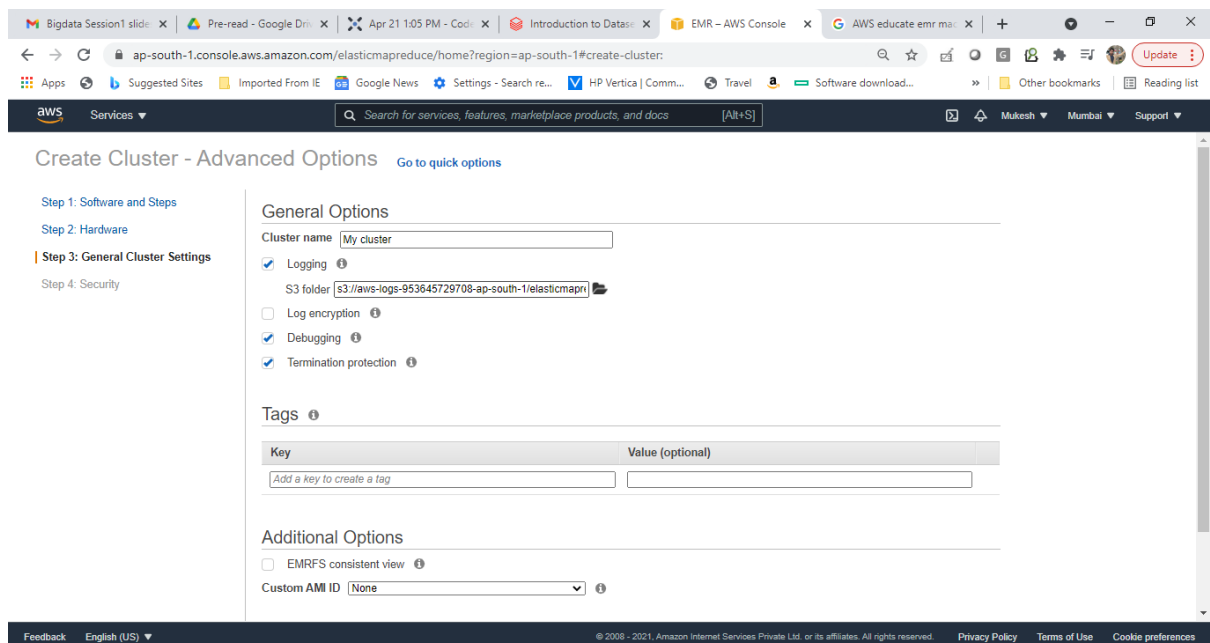
Next

5.

The screenshot shows the AWS EMR console interface for creating a cluster. The browser tabs include 'Bigdata Session1 slide', 'Pre-read - Google Drive', 'Apr 21 1:05 PM - Code', 'Introduction to DataS...', 'EMR - AWS Console', and 'AWS educate emr ma...'. The address bar shows the URL: `ap-south-1.console.aws.amazon.com/elasticmapreduce/home?region=ap-south-1#create-cluster`. The page title is 'Create Cluster - Advanced Options' with a link to 'Go to quick options'. The left sidebar shows the steps: 'Step 1: Software and Steps', 'Step 2: Hardware' (selected), 'Step 3: General Cluster Settings', and 'Step 4: Security'. The main content area is divided into 'Hardware Configuration' and 'Cluster Composition'. Under 'Cluster Composition', the 'Instance group configuration' section has two options: 'Uniform instance groups' (selected) and 'Instance fleets'. The 'Networking' section is partially visible at the bottom. The footer includes 'Feedback', 'English (US)', '© 2018 - 2021, Amazon Internet Services Private Ltd. or its affiliates. All rights reserved.', 'Privacy Policy', 'Terms of Use', and 'Cookie preferences'.



6. Now click on the “Next” button



7. Next screen is “Security Options” screen . here in this screen , you need to select “EC Key pair” which we created in the 1st step

Bigdata Session1 slide X Pre-read - Google Dr X Apr 21 1:05 PM - Cod X Introduction to Datas X EMR - AWS Console X AWS educate emr ma X +

ap-south-1.console.aws.amazon.com/elasticmapreduce/home?region=ap-south-1#create-cluster-j-1JEZWBXNZ1WMR

Services Search for services, features, marketplace products, and docs [Alt+S]

Create Cluster - Advanced Options [Go to quick options](#)

Step 1: Software and Steps
Step 2: Hardware
Step 3: General Cluster Settings
Step 4: Security

Security Options

EC2 key pair **emrkeypair**

☒ Cluster visible to all IAM users in account

Permissions

☒ Default ☐ Custom
Use default IAM roles. If roles are not present, they will be automatically created for you with managed policies for automatic policy updates.

EMR role [EMR_DefaultRole](#)

EC2 instance profile [EMR_EC2_DefaultRole](#)

Auto Scaling role [EMR_AutoScaling_DefaultRole](#)

Security Configuration

Security configuration **None**

EC2 security groups

An EC2 security group acts as a virtual firewall for your cluster nodes to control inbound and outbound traffic. There are two types of security groups you can configure, [EMR managed security groups](#) and [additional security groups](#). EMR will [automatically update](#) the rules in the EMR managed security groups in order to launch a cluster. [Learn more](#)

Feedback English (US) © 2008 - 2021, Amazon Internet Services Private Ltd. or its affiliates. All rights reserved. Privacy Policy Terms of Use Cookie preferences

emrkeypair.ppk Show all

8.

Bigdata Session1 slide X Pre-read - Google Dr X Apr 21 1:05 PM - Cod X Introduction to Datas X EMR - AWS Console X AWS educate emr ma X +

ap-south-1.console.aws.amazon.com/elasticmapreduce/home?region=ap-south-1#cluster-details-j-1JEZWBXNZ1WMR

Services Search for services, features, marketplace products, and docs [Alt+S]

Amazon EMR

EMR Studio
EMR on EC2
Clusters
Notebooks
Git repositories
Security configurations
Block public access
VPC subnets
Events
EMR on EKS
Virtual clusters
Help
What's new

Clone Terminate AWS CLI export

Cluster: My cluster Starting

Summary Application user interfaces Monitoring Hardware Configurations Events Steps Bootstrap actions

Summary

ID: j-1JEZWBXNZ1WMR
Creation date: 2021-04-24 22:34 (UTC+5:30)
Elapsed time: 0 seconds
After last step completes: Cluster waits
Termination protection: On [Change](#)
Tags: -- [View All / Edit](#)
Master public DNS: --

Configuration details

Release label: emr-5.33.0
Hadoop distribution: Amazon 2.10.1
Applications: Hive 2.3.7, Hue 4.9.0, Spark 2.4.7, Sqoop 1.4.7
Log URI: s3://aws-logs-953645729708-ap-south-1/elasticmapreduce/
EMRFS consistent view: Disabled
Custom AMI ID: --

Application user interfaces

Persistent user interfaces: --
On-cluster user interfaces: --

Network and hardware

Availability zone: --
Subnet ID: [subnet-8ad15bf1](#)
Master: **Provisioning** 1 c4.xlarge
Core: **Provisioning** 1 c4.xlarge
Task: --
Cluster scaling: Not enabled

Security and access

Key name: 14dec2020
EC2 instance profile: EMR_EC2_DefaultRole
EMR role: EMR_DefaultRole

Feedback English (US) © 2008 - 2021, Amazon Internet Services Private Ltd. or its affiliates. All rights reserved. Privacy Policy Terms of Use Cookie preferences