

Cloud Computing Week 9

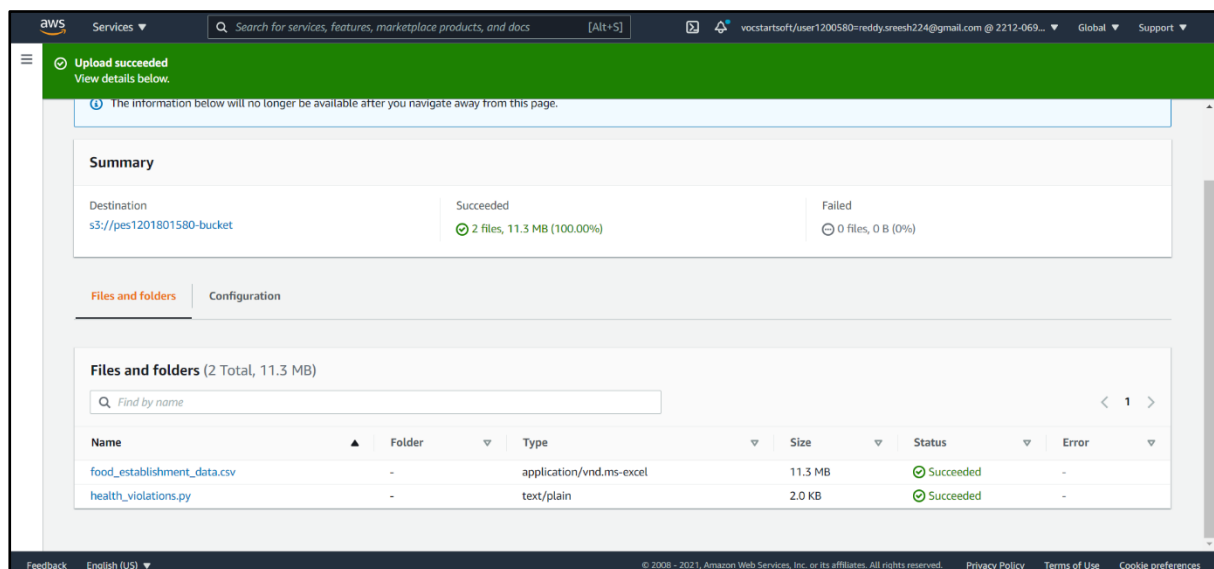
Name: K. SREESH REDDY

SRN: PES1201801580

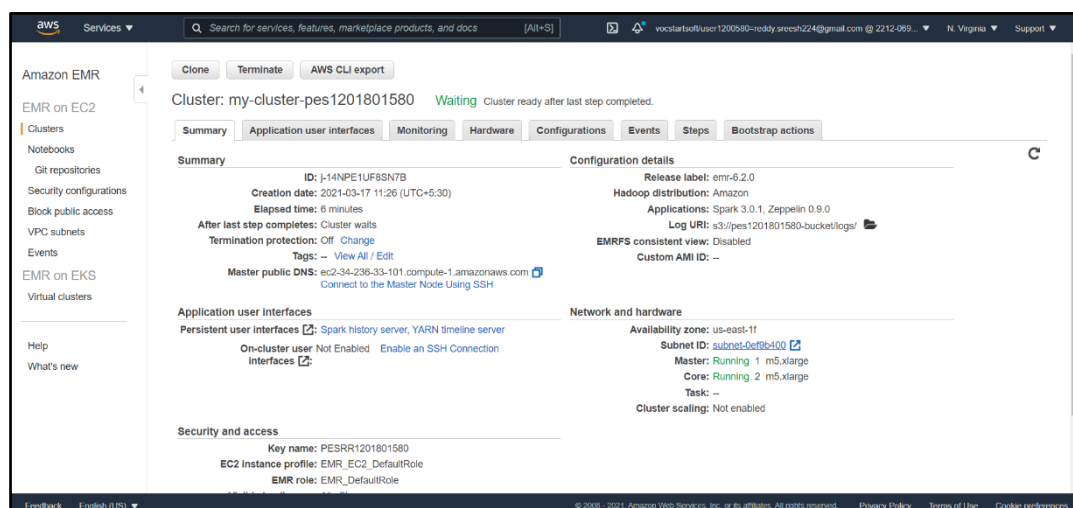
Section: E

TASK 1: Plan and Configure an Amazon EMR Cluster

Upload both .csv and .py files into S3 bucket and take a screenshot with uploaded files required (9a)



Screenshot of running cluster (9b)



TASK2: Managing Amazon EMR Clusters

Screenshot of spark job completed (9c)

Cluster: my-cluster-pes1201801580 Waiting Cluster ready after last step completed.

Summary Application user interfaces Monitoring Hardware Configurations **Steps** Bootstrap actions

Concurrency: 1 [Change](#)
After last step completes: Cluster waits

[Add step](#) [Clone step](#) [Cancel step](#)

Filter: All steps Filter steps ... 2 steps (all loaded) [View Jobs in the Application History Tab](#)

	ID	Name	Status	Start time (UTC+5:30)	Elapsed time	Log files
	s-BQH96WQFVP0Z	Spark application	Completed	2021-03-17 11:38 (UTC+5:30)	36 seconds	View logs
	s-3FKIPQDGPXMS4	Setup hadoop debugging	Completed	2021-03-17 11:32 (UTC+5:30)	6 seconds	View logs

Screenshot of output folder in S3 (9d)

Amazon S3 > pes1201801580-bucket > my-output/

my-output/ [Copy S3 URI](#)

Objects Properties

Objects (2)
Objects are the fundamental entities stored in Amazon S3. You can use [Amazon S3 inventory](#) to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. [Learn more](#)

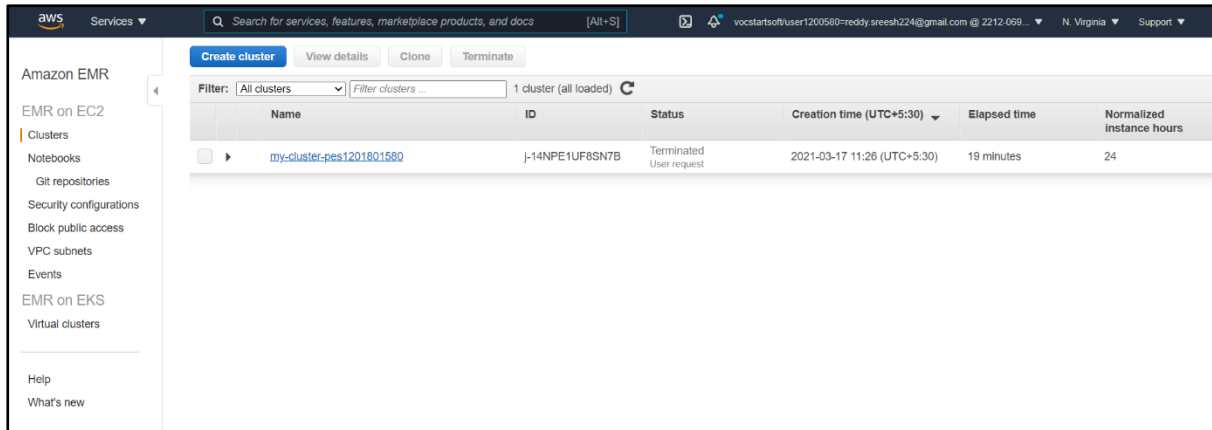
[Refresh](#) [Delete](#) [Actions](#) [Create folder](#) [Upload](#)

Find objects by prefix

	Name	Type	Last modified	Size	Storage class
<input type="checkbox"/>	_SUCCESS	-	March 17, 2021, 11:38:40 (UTC+05:30)	0 B	Standard
<input type="checkbox"/>	part-00000-e670f199-d7e3-4550-8142-208570a51a62-c000.csv	csv	March 17, 2021, 11:38:40 (UTC+05:30)	219.0 B	Standard

TASK3: Clean Up Amazon EMR Cluster Resources

Go to EMR, select your cluster and click “Terminate”



The screenshot shows the AWS Management Console interface for Amazon EMR. The left sidebar contains navigation links for Amazon EMR, EMR on EC2, Clusters, Notebooks, Git repositories, Security configurations, Block public access, VPC subnets, Events, EMR on EKS, and Virtual clusters. The main content area displays a table of EMR clusters. The table has columns for Name, ID, Status, Creation time (UTC+5:30), Elapsed time, and Normalized instance hours. A single cluster is listed with the name 'my-cluster-pes1201801580', ID 'j-14NPE1UF8SN7B', and status 'Terminated (User request)'. The creation time is '2021-03-17 11:26 (UTC+5:30)', the elapsed time is '19 minutes', and the normalized instance hours are '24'.

Name	ID	Status	Creation time (UTC+5:30)	Elapsed time	Normalized instance hours
my-cluster-pes1201801580	j-14NPE1UF8SN7B	Terminated User request	2021-03-17 11:26 (UTC+5:30)	19 minutes	24

Csv file name: part-00000-e670f199-d7e3-4550-8142-208570a51a62-c000