

Cloud Computing Lab 9 - AWS EMR

PES1201800177

Abhishek Das

Section : G

9a

Amazon S3 > pes1201800177-bucket

pes1201800177-bucket

Objects | Properties | Permissions | Metrics | Management | Access Points

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

<input type="checkbox"/>	Name	Type	Last modified	Size	Storage class
<input type="checkbox"/>	food_establishment_data.csv	csv	March 22, 2021, 18:35:43 (UTC+05:30)	11.3 MB	Standard
<input type="checkbox"/>	health_violations.py	py	March 22, 2021, 18:35:44 (UTC+05:30)	2.0 KB	Standard

9b

[Clone](#) [Terminate](#) [AWS CLI export](#)

Cluster: My-cluster-PES1201800177 **Starting**

[Summary](#) [Application user interfaces](#) [Monitoring](#) [Hardware](#) [Configurations](#) [Events](#) [Steps](#) [Bootstrap actions](#)

Summary

ID: j-2UYG1JALKULPP
Creation date: 2021-03-22 18:39 (UTC+5:30)
Elapsed time: 1 second
After last step completes: Cluster waits
Termination protection: Off [Change](#)
Tags: -- [View All](#) / [Edit](#)
Master public DNS: --

Configuration details

Release label: emr-6.2.0
Hadoop distribution: Amazon
Applications: Spark 3.0.1, Zeppelin 0.9.0
Log URI: s3://pes1201800177-bucket/logs/ [View](#)
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-4cf9686d](#)
Master: [Provisioning](#) 1 m5.xlarge
Core: [Provisioning](#) 2 m5.xlarge
Task: --
Cluster scaling: Not enabled

Security and access

Key name: lab1
EC2 instance profile: EMR_EC2_DefaultRole
EMR role: EMR_DefaultRole
Visible to all users: All [Change](#)
Security groups for Master:
Security groups for Core & Task:

9c

Clone

Terminate

AWS CLI export

Cluster: My-cluster-PES1201800177 Running Running step

Summary

Application user interfaces

Monitoring

Hardware

Configurations

Events

Steps

Bootstrap actions

Concurrency: 1 [Change](#)

After last step completes: Cluster waits

Add step

Clone step

Cancel step

Filter: All steps 2 steps (all loaded)

	ID	Name	Status	Start time (UTC+5:30)	Elapsed time	Log files
	s-BFFSVBKEU0HG	Spark application	Running	2021-03-22 18:46 (UTC+5:30)	1 minute	View logs
<div>JAR location : command-runner.jar</div> <div>Main class : None</div> <div>Arguments : spark-submit --deploy-mode cluster s3://pes1201800177-bucket/health_violations.py --data_source s3://pes1201800177-bucket/food_establishment_data.csv --output_uri s3://pes1201800177-bucket/health_violations_data.csv</div> <div>Action on failure: Continue</div>						
	s-VANKIT5TNVUC	Setup hadoop debugging	Completed	2021-03-22 18:45 (UTC+5:30)	4 seconds	View logs

9d

Amazon S3 > pes1201800177-bucket > myOutputFolder/

myOutputFolder/

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](#)

Delete

Actions

Create folder

Upload

<input type="checkbox"/>	Name	Type	Last modified	Size
<input type="checkbox"/>	_SUCCESS	-	March 22, 2021, 18:46:36 (UTC+05:30)	
<input type="checkbox"/>	part-00000-56e09bb4-bac4-4faf-9f7d-89a1829c6671-c000.csv	csv	March 22, 2021, 18:46:36 (UTC+05:30)	