

Name








lab1-cluster

Amazon EMR release [Info](#)

A release contains a set of applications which can be installed on your cluster.

emr-7.3.0

Application bundle

Spark Interactive 	Core Hadoop 	Flink 	HBase 	Presto 	Trino 	Custom 
---	---	--	--	---	---	---

- |   |  |  |
|---|--|--|
| <input type="checkbox"/> AmazonCloudWatchAgent 1.300032.2 | <input type="checkbox"/> Flink 1.18.1                              | <input type="checkbox"/> HBase 2.4.17                |
| <input checked="" type="checkbox"/> HCatalog 3.1.3        | <input checked="" type="checkbox"/> Hadoop 3.3.6                   | <input checked="" type="checkbox"/> Hive 3.1.3       |
| <input checked="" type="checkbox"/> Hue 4.11.0            | <input checked="" type="checkbox"/> JupyterEnterpriseGateway 2.6.0 | <input checked="" type="checkbox"/> JupyterHub 1.5.0 |
| <input checked="" type="checkbox"/> Livy 0.8.0            | <input type="checkbox"/> Oozie 5.2.1                               | <input type="checkbox"/> Phoenix 5.1.3               |
| <input type="checkbox"/> Pig 0.17.0                       | <input type="checkbox"/> Presto 0.285                              | <input checked="" type="checkbox"/> Spark 3.5.1      |
| <input checked="" type="checkbox"/> Sqoop 1.4.7           | <input type="checkbox"/> TensorFlow 2.16.1                         | <input type="checkbox"/> Tez 0.10.2                  |
| <input type="checkbox"/> Trino 442                        | <input checked="" type="checkbox"/> Zeppelin 0.11.1                | <input checked="" type="checkbox"/> ZooKeeper 3.9.1  |

AWS Glue Data Catalog settings

Use the AWS Glue Data Catalog to provide an external metastore for your application.

- ☒ Use for Hive table metadata
- ☒ Use for Spark table metadata

Operating system options [Info](#)

- ☒ Amazon Linux release
  - ☐ Custom Amazon Machine Image (AMI)
- ☒ Automatically apply latest Amazon Linux updates

## ▼ Cluster configuration - required [Info](#)

Choose a configuration method for the primary, core, and task node groups for your cluster.

### ☒ Uniform instance groups

Choose the same EC2 instance type and purchasing option (On-Demand or Spot) for all nodes in your node group. [Learn more](#)

### ☐ Flexible instance fleets

Choose from the widest variety of provisioning options for the EC2 instances in your cluster. Diversify instance types and purchasing options, and use an allocation strategy. [Learn more](#)

## Uniform instance groups

### Primary

Choose EC2 instance type

m5.xlarge  
4 vCore 16 GiB memory  
EBS only storage On-Demand price: -  
Lowest Spot price: -

Actions ▼

### ☐ Use high availability

Launch highly available, more resilient cluster with three primary nodes on On-Demand Instances. This configuration applies for the lifetime of your cluster. [Learn more](#)

### ► Node configuration - optional

### Core

Choose EC2 instance type

m5.xlarge  
4 vCore 16 GiB memory  
EBS only storage On-Demand price: -  
Lowest Spot price: -

Actions ▼

### ► Node configuration - optional

▼ **Software settings** [Info](#)

Override the default configurations for specific applications on your cluster.

☒ Enter configuration

☐ Load JSON from Amazon S3

1▼ [ {  
2▼ {  
3 "Classification": "jupyter-s3-conf",  
4▼ "Properties": {  
5 "s3.persistence.enabled": "true",  
6 "s3.persistence.bucket": "labs-bucket-telematica"  
7 }  
8 }  
9 ]

JSON Ln 9, Col 2 ✖ : 0 ⚠ : 0 ⚙

vockey

×

Browse

Create key pair [↗](#)

▼ **Identity and Access Management (IAM) roles - required** [Info](#)

Choose or create a service role and instance profile for the EC2 instances in your cluster.

**Amazon EMR service role** [Info](#)

The service role is an IAM role that Amazon EMR assumes to provision resources and perform service-level actions with other AWS services.

☒ Choose an existing service role

Select a default service role or a custom role with IAM policies attached so that your cluster can interact with other AWS services.

☐ Create a service role

Let Amazon EMR create a new service role so that you can grant and restrict access to resources in other AWS services.

Service role

EMR\_DefaultRole ▼

↻

**EC2 instance profile for Amazon EMR**

The instance profile assigns a role to every EC2 instance in a cluster. The instance profile must specify a role that can access the resources for your steps and bootstrap actions.

☒ Choose an existing instance profile

Select a default role or a custom instance profile with IAM policies attached so that your cluster can interact with your resources in Amazon S3.

☐ Create an instance profile

Let Amazon EMR create a new instance profile so that you can specify a custom set of resources for it to access in Amazon S3.

Instance profile

EMR\_EC2\_DefaultRole ▼

↻

**Custom automatic scaling role - optional**

When a custom automatic scaling rule triggers, Amazon EMR assumes this role to add and terminate EC2 instances. [Learn more](#) [↗](#)

Custom automatic scaling role

EMR\_AutoScaling\_DefaultRole ▼

↻

Create IAM role [↗](#)