Clone "Retail Online" Info

Name and applications Info

Name
Retail Online

Amazon EMR release Info
A release contains a set of applications which can be installed on your cluster.
emr-5.30.0

Application bundle
Spark | Core Hadoop | HBase | Presto | Custom

- [ ] Flink 1.10.0
- [ ] HCatalog 2.3.6
- [ ] Hue 4.6.0
- [ ] MXNet 1.5.1
- [ ] Phoenix 4.14.3
- [x] Spark 2.4.5
- [ ] Tez 0.9.2
- [ ] Ganglia 3.7.2
- [x] Hadoop 2.8.5
- [ ] JupyterHub 1.1.0
- [ ] Mahout 0.13.0
- [ ] Pig 0.17.0
- [ ] Sqoop 1.4.7
- [x] Zeppelin 0.8.2
- [ ] HBase 1.4.13
- [ ] Hive 2.3.6
- [ ] Livy 0.7.0
- [ ] Oozie 5.2.0
- [ ] Presto 0.232
- [ ] TensorFlow 1.14.0
- [ ] ZooKeeper 3.4.14

AWS Glue Data Catalog settings
Use the AWS Glue Data Catalog to provide an external metastore for your application.
- [ ] Use for Spark table metadata

Custom Amazon Machine Image (AMI) Info
Choose or enter an AMI ID

- [x] Update all installed packages on reboot

Summary Info

Name and applications
Name
Retail Online

Amazon EMR release
emr-5.30.0

Application bundle
Custom (Hadoop 2.8.5, Spark 2.4.5, Zeppelin 0.8.2)

Cluster configuration
Instance groups
Primary (m4.xlarge)

Networking
VPC
vpc-0136574cf...

Subnet
subnet-0f95a8...

Cancel | Clone cluster

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

- (•) Instance groups — Choose one instance type per node group
- ( ) Instance fleets — Choose any combination of instance types within each node group

Instance groups

Primary
Choose EC2 instance type
m4.xlarge
4 vCore 16 GiB memory EBS only storage
On-Demand price: - Lowest Spot price: -
Actions

- [ ] Use multiple primary nodes
To improve cluster availability, use 3 primary nodes with the same configuration and bootstrap actions. You can not use multiple primary nodes with instance fleets.

▶ Node configuration - *optional*

Add core instance group
You must add a core instance group before you can add a task instance group.

EBS root volume
EBS root volume applies to the operating systems and applications that you install on the cluster.
Size (GiB)
40
10 - 100 GiB per volume General Purpose SSD (gp2)

Networking Info
Virtual private cloud (VPC) Info
vpc-0136574cf6db99ef7 Browse Create VPC

Subnet Info
subnet-0f95a865a5f684daf Browse Create subnet

▶ EC2 security groups (firewall)

▶ Steps - *optional* (0) Info
Remove Edit Add
Use commands and scripts to tell your cluster where to find and how to process your data. Steps run consecutively unless you enable the Concurrency option.

Cluster termination Info
- (•) Manually terminate cluster
- ( ) Automatically terminate cluster after last step ends
- ( ) Automatically terminate cluster after idle time (Recommended)

- [x] Use termination protection
Protect your EC2 instances from accidental termination.

▶ Bootstrap actions - *optional* Info
Use bootstrap actions to install software or customize your instance configuration.

▶ Cluster logs - *optional* Info

▶ Tags - *optional* Info
A tag is a label that you assign to an AWS resource. Each tag consists of a key and an optional value. You can use tags to search and filter your resources or track your AWS costs.

▶ Software settings - *optional* Info

Security configuration and EC2 key pair - *optional* Info

Security configuration
Select your cluster encryption, authentication, authorization, and instance metadata service settings.
Choose a security configuration Browse Create security configuration

Amazon EC2 key pair for SSH to the cluster Info
debskp Browse Create key pair

Identity and Access Management (IAM) roles 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
Choose IAM role Create IAM role

Debdatta Sarkar

© 2023, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences