

Service Summary Cards (SSC)

Reference:

[FAQs](#)

Category:

Analytics



Amazon EMR

What?

- Amazon EMR is cloud big data platform for data processing, interactive analysis, and machine learning using open source frameworks such as Apache Spark, Apache Hive, and Presto.

Why?

- Amazon EMR simplifies building and operating big data environments and applications.
- Amazon EMR enables you to quickly and easily provision as much capacity as you need, and automatically or manually add and remove capacity. This is very useful if you have variable or unpredictable processing requirements.

When?

- When you want to focus on transforming and analyzing your data without having to worry about infrastructure provisioning, cluster setup, configuration, open-source applications or tuning.

Where?

- Amazon EMR launches all nodes (Master, Core, Optional Task nodes) for a given cluster in the same Availability Zone of a Region.

Who?

- After an EMR cluster is launched, customer can monitor and manage it. Amazon EMR provides several tools you can use to connect to and control your cluster.

How?

- You can launch a cluster by specifying the name of your cluster, the location in Amazon S3 of your input data, your processing application, your desired data output location, and the number and type of Amazon EC2 instances you'd like to use.

How much?

- You pay a per-second rate for every second you use, with a one-minute minimum. Customers pay for Amazon EMR plus backend compute price (Amazon EC2, Amazon EKS, AWS Outposts, Amazon EMR Serverless).

More SSCs:

[Click Here](#)

Complete Book

[Click Here](#)

Created by:

[Ashish Prajapati](#)

