

Managed EMR

1. 소프트웨어 및 단계

1. 릴리즈 : **emr-6.0.0**

어플리케이션 : **Hadoop 3.2.1, Hive 3.1.2, Hue 4.4.0**

클러스터 생성 - 고급 옵션 [백문 옵션으로 이동](#)

단계 1: 소프트웨어 및 단계

단계 2: 하드웨어

단계 3: 일반 클러스터 설정

단계 4: 보안

소프트웨어 구성

릴리스 **emr-6.0.0**

<input checked="" type="checkbox"/> Hadoop 3.2.1	<input type="checkbox"/> Zeppelin 0.9.0	<input type="checkbox"/> Livy 0.6.0
<input type="checkbox"/> JupyterHub 1.0.0	<input type="checkbox"/> Tez 0.9.2	<input type="checkbox"/> Ganglia 3.7.2
<input type="checkbox"/> HBase 2.2.3	<input checked="" type="checkbox"/> Hive 3.1.2	<input type="checkbox"/> Presto 0.230
<input type="checkbox"/> ZooKeeper 3.4.14	<input type="checkbox"/> MXNet 1.5.1	<input checked="" type="checkbox"/> Hue 4.4.0
<input type="checkbox"/> Phoenix 5.0.0	<input type="checkbox"/> Oozie 5.1.0	<input type="checkbox"/> Spark 2.4.4
<input type="checkbox"/> HCatalog 3.1.2	<input type="checkbox"/> TensorFlow 1.14.0	

☒ Pig은(는) 선택한 릴리스에서 지원되지 않습니다.

멀티 마스터 지원

☐ 여러 마스터 노드를 사용하여 클러스터 가용성을 개선합니다. [자세히 알아보기](#)

AWS Glue 데이터 카탈로그 설정(선택 사항)

☐ Hive 테이블 메타 데이터에서 사용

소프트웨어 설정 편집

☒ 구성 입력 ☐ S3에서 JSON 로드

`classification=config-file-name.properties[myKey1=myValue1,myKey2=myValue2]`

모든 단계(선택 사항)

단계는 클러스터에 제출하는 작업의 단위입니다. 예를 들어, 한 단계에 하나 이상의 Hadoop 또는 Spark 작업이 포함될 수 있습니다. 실행 후 추가 단계를 클러스터에 제출할 수도 있습니다. [자세히 알아보기](#)

Concurrency: ☐ Run multiple steps at the same time to improve cluster utilization

After last step completes: ☒ Clusters enters waiting state

☐ 마지막 단계가 완료된 후 클러스터 자동 종료

단계 유형 **[단계 선택]**

단계 추가

2. 하드웨어

1. Cluster Composition : **균일한 인스턴스 그룹**

2. Networking : **public subnet**

클러스터 생성 - 고급 옵션 [백문 옵션으로 이동](#)

단계 1: 소프트웨어 및 단계

단계 2: 하드웨어

단계 3: 일반 클러스터 설정

단계 4: 보안

하드웨어 구성

Specify the networking and hardware configuration for your cluster. Request Spot instances (unused EC2 capacity) to save money.

Cluster Composition

EC2 인스턴스가 20개 이상 필요한 경우 [이 주제를 참조하십시오](#)

인스턴스 그룹 구성

☒ **균일한 인스턴스 그룹**
각 노드 유형별로 단일 인스턴스 유형 및 구매 옵션을 지정합니다.

☐ 인스턴스 플릿
목표 용량과 Amazon EMR에서 각 노드 유형별로 이 목표를 달성할 수 있는 방법을 지정합니다. 최소 인스턴스 유형과 구매 옵션입니다. [자세히 알아보기](#)

Networking

Use a Virtual Private Cloud (VPC) to process sensitive data or connect to a private network. Launch the cluster into a VPC with a public, private or shared subnet. Subnets may be associated with and AWS Outpost or AWS Local Zone.

Launch the cluster into a VPC with a public, private, or shared subnet. Subnets may be associated with an AWS Outpost or AWS Local Zone.

네트워크 **vpc-02067642b982c2cf7 | cjm-vpc** [VPC 만들기](#)

EC2 서브넷 **subnet-08f17a12a46396972 | cjm-subnet-public-a | us-west-2a**

3. Cluster Nodes and Instances

- master : **m5.xlarge, 1개**
- core : **m5.xlarge, 1개**
- task : **m5.xlarge, 2개**

4. Cluster scaling : **disable**

5. EBS Root Volume : **100 GB**

Cluster Nodes and Instances

인스턴스 유형, 인스턴스 수 및 구매 옵션을 선택합니다. 온디맨드 인스턴스, 스팟 인스턴스 또는 둘 다 사용하도록 선택할 수 있습니다. 인스턴스 유형 및 구매 옵션이 각 인스턴스 그룹의 모든 EC2 인스턴스에 적용되므로 인스턴스 그룹 생성 시 해당 인스턴스 그룹에 대해서만 이러한 옵션을 지정하면 됩니다. [인스턴스 구매 옵션에 대해 자세히 알아보기](#)

Console options for automatic scaling have changed. [Learn more](#)

노드 유형	인스턴스 유형	인스턴스 개수	구매 옵션
마스터 마스터 인스턴스 그룹 - 1	m5.xlarge 4개의 vCore, 16 GiB 메모리, EBS 전용 스토리지 EBS 스토리지: 64GiB 구성 설정 추가	1 인스턴스	<input checked="" type="radio"/> 온디맨드 <input type="radio"/> 스팟 온디맨드를 최대 가격으로 사용
코어 코어 인스턴스 그룹 - 2	m5.xlarge 4개의 vCore, 16 GiB 메모리, EBS 전용 스토리지 EBS 스토리지: 64GiB 구성 설정 추가	<input type="text" value="3"/> 인스턴스	<input checked="" type="radio"/> 온디맨드 <input type="radio"/> 스팟 온디맨드를 최대 가격으로 사용

+ 작업 인스턴스 그룹 추가

Total core and task units

3 총 유닛

Cluster scaling

Adjust the number of Amazon EC2 instances available to an EMR cluster via EMR-managed scaling or a custom automatic scaling policy. [Learn more](#)

Cluster scaling ☐ Enable Cluster Scaling

EBS Root Volume

Specify the root device volume size up to 100 GiB. This sizing applies to all instances in the cluster. [Learn more](#)

루트 디바이스 EBS 볼륨 크기 GiB

3. 일반 클러스터 설정

1. 클러스터 이름 : **cjm_emr**

2. 디버깅 : **enable**

3. 종료보호 : **disable**

클러스터 생성 - 고급 옵션 [백린 옵션으로 이동](#)

단계 1: 소프트웨어 및 단계

단계 2: 하드웨어

단계 3: 일반 클러스터 설정

단계 4: 보안

일반 옵션

클러스터 이름

☒ 로깅
S3 볼더

☒ 디버깅
☐ 종료 보호

태그

키	값(선택 사항)
Name	cjm_emr_cluster
Owner	cjm
expiry-date	2021-06-10
키를 추가하여 태그 생성	

추가 옵션

☐ EMRFS 일관된 보기
사용자 지정 AMI ID

부트스트랩 작업

취소

이전

다음

4. 보안

1. EC2 키 페어 : {본인 key pair}

2. EMR Role, EC2 Instance profile, Auto Scaling Role : **default**

클러스터 생성 - 고급 옵션 [백문 옵션으로 이동](#)

[단계 1: 소프트웨어 및 단계](#)

[단계 2: 하드웨어](#)

[단계 3: 일반 클러스터 설정](#)

단계 4: 보안

보안 옵션

EC2 키 페어 [cjm-key](#) ⓘ

☒ 계정에 있는 모든 IAM 사용자가 볼 수 있는 클러스터 ⓘ

권한 ⓘ

☒ 기본값 ☐ 사용자 지정

기본 IAM 역할 사용. 역할이 없으면 자동 정책 업데이트에 대한 관리형 정책을 사용하여 자동으로 생성됩니다.

EMR 역할 [EMR_DefaultRole](#) ⓘ

EC2 인스턴스 프로파일 [EMR_EC2_DefaultRole](#) ⓘ

Auto Scaling 역할 [EMR_AutoScaling_DefaultRole](#) ⓘ

▶ 보안 구성

▼ EC2 보안 그룹

EC2 보안 그룹은 클러스터 노드에 대한 인바운드 및 아웃바운드 트래픽을 제어하는 가상 방화벽 역할을 합니다. [EMR 관리형 보안 그룹](#)과 [추가 보안 그룹](#)이라는 두 가지 유형의 보안 그룹을 구성할 수 있습니다. EMR은 클러스터를 시작하기 위해 EMR 관리형 보안 그룹에서 규칙을 [자동으로 업데이트](#)합니다. [자세히 알아보십시오](#)

유형	EMR 관리형 보안 그룹 EMR은 선택한 그룹을 기본으로 설정 합니다.	추가 보안 그룹 EMR은 선택한 그룹을 수정하지 않습니다.
마스터	기본값 : sg-0c03f241758cdd591 (ElasticMapReduce) 선택한 보안 그룹 없음	
코어 및 작업	기본값 : sg-0fb57e6947d382104 (ElasticMapReduce) 선택한 보안 그룹 없음	

[보안 그룹 설정](#)

[취소](#)

[이전](#)

[클러스터 생성](#)