Create Dataproc cluster

Select the infrastructure service that you want to use.

Cluster on Compute Engine
Create the cluster on Compute Engine.
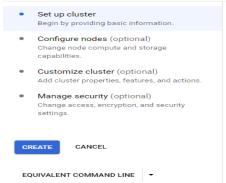
`CREATE`
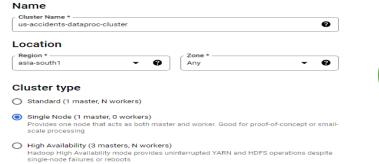
Cluster on GKE
Create the cluster on Google Kubernetes Engine (GKE).

`CREATE`

2

CANCEL

← Create a Dataproc cluster on Compute Engine

- **Set up cluster**
  Begin by providing basic information.
- **Configure nodes** (optional)
  Change node compute and storage capabilities.
- **Customize cluster** (optional)
  Add cluster properties, features, and actions.
- **Manage security** (optional)
  Change access, encryption, and security settings.

`CREATE` CANCEL

EQUIVALENT COMMAND LINE ▼

Name

Cluster Name *
us-accidents-dataproc-cluster ❷

Location

Region *
asia-south1 ❷

Zone *
Any ❷

Cluster type

○ Standard (1 master, N workers)

◉ Single Node (1 master, 0 workers)
Provides one node that acts as both master and worker. Good for proof-of-concept or small-scale processing

○ High Availability (3 masters, N workers)
Hadoop High Availability mode provides uninterrupted YARN and HDFS operations despite single-node failures or reboots

3

← Create a Dataproc cluster on Compute Engine

- Set up cluster
  Begin by providing basic information.
- **Configure nodes** (optional)
  Change node compute and storage capabilities.
- Customize cluster (optional)
  Add cluster properties, features, and actions.
- Manage security (optional)
  Change access, encryption, and security settings.

`CREATE` CANCEL

EQUIVALENT COMMAND LINE ▼

✓ General purpose

Machine types for common workloads, optimized for cost and flexibility

Series
N2 ▼

Powered by Intel Cascade Lake and Ice Lake CPU platforms

Machine type
n2-standard-4 (4 vCPU, 2 core, 16 GB memory) ▼

	vCPU	Memory
	4	3.75 GB

∨ CPU PLATFORM AND GPU

Primary disk size *
30 GB ❷

Primary disk type *
Balanced Persistent Disk ▼ ❷

Number of local SSDs *
0 ▼ x 375GB ❷

Local SSD Interface
SCSI ▼ ❷

4

- Set up cluster
  Begin by providing basic information.
- Configure nodes (optional)
  Change node compute and storage capabilities.
- **Customize cluster** (optional)
  Add cluster properties, features, and actions.
- Manage security (optional)
  Change access, encryption, and security settings.

`CREATE` CANCEL

Internal IP only

☐ Configure all instances to have only internal IP addresses. Learn more ↗

Labels

A list of key:value pairs to attach to the cluster for tracking.

+ ADD LABELS

Cluster properties

Use cluster properties to add or modify configuration files when creating a cluster.

+ ADD PROPERTIES

Initialization actions

5