



Module 2

Big Data Processing using Cloudera Quickstart with a Docker Container

Thanachart Numnonda, Executive Director, IMC Institute

Thanisa Numnonda, Faculty of Information Technology,
King Mongkut's Institute of Technology Ladkrabang

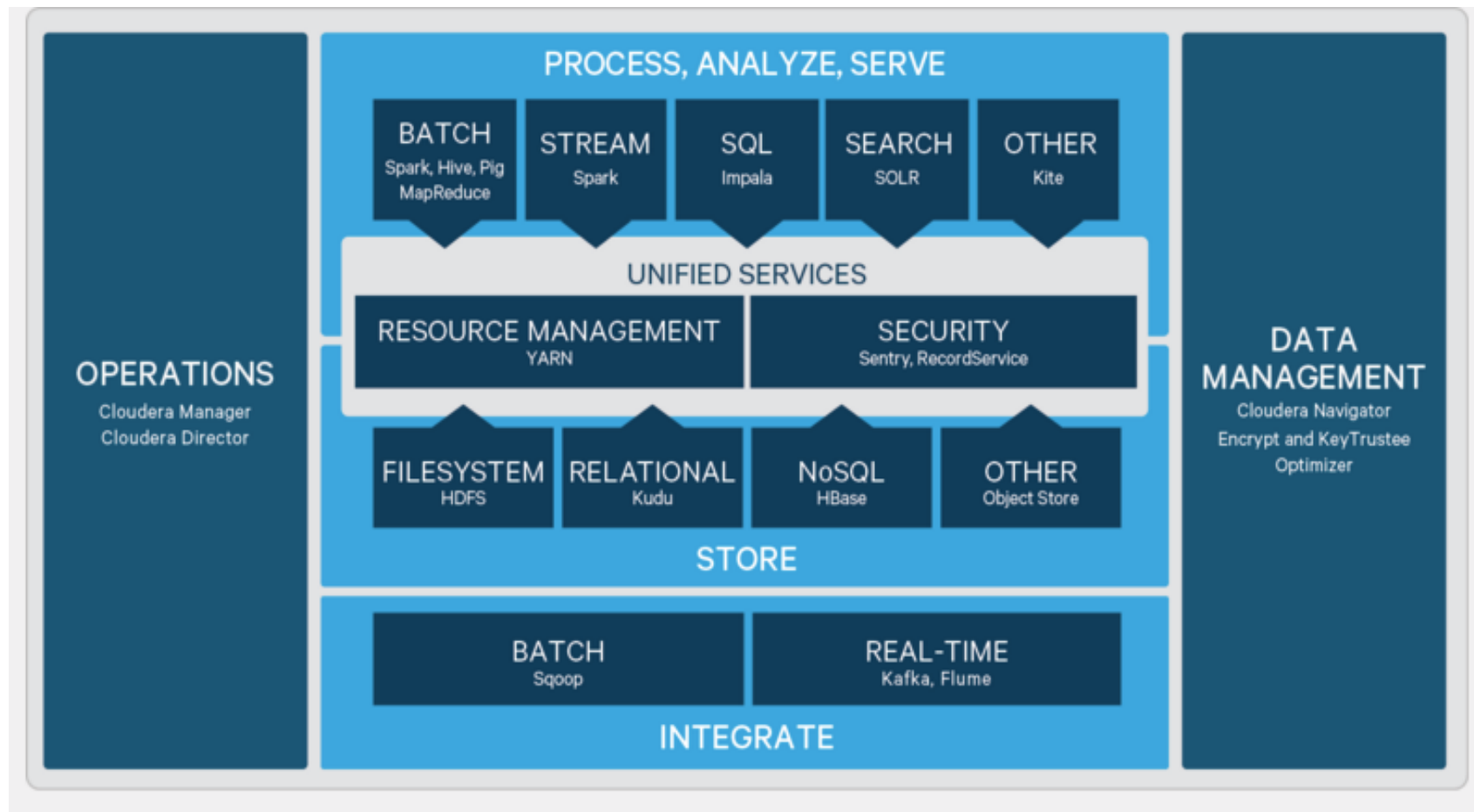


Hadoop Distribution

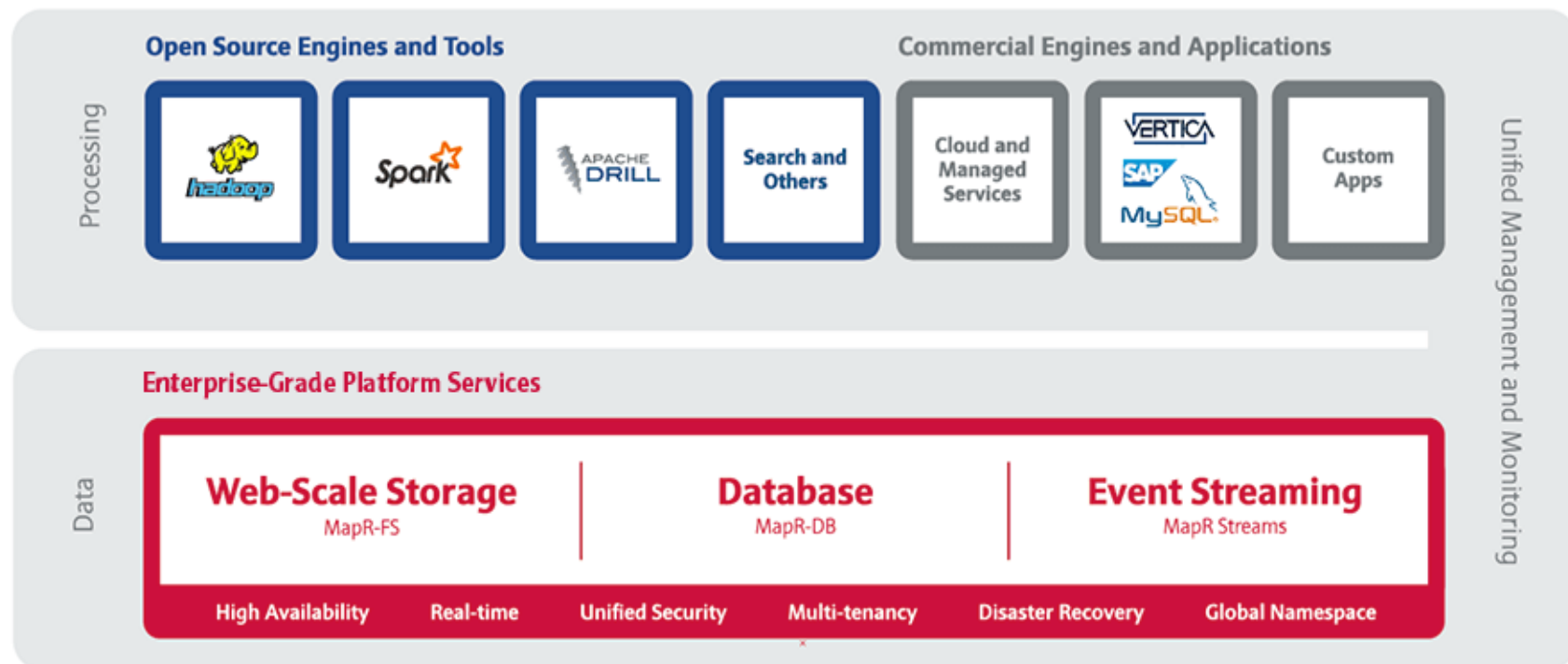
- **On-Premise**
 - Pure Apache Hadoop
 - Cloudera
 - MapR
 - Hortonworks
 - Pivotal HDB
 - IBM Infosphere BigInsight
- **On-Cloud (Hadoop as a Service)**
 - Amazon EMR
 - Microsoft Azure HDInsight
 - Google Cloud



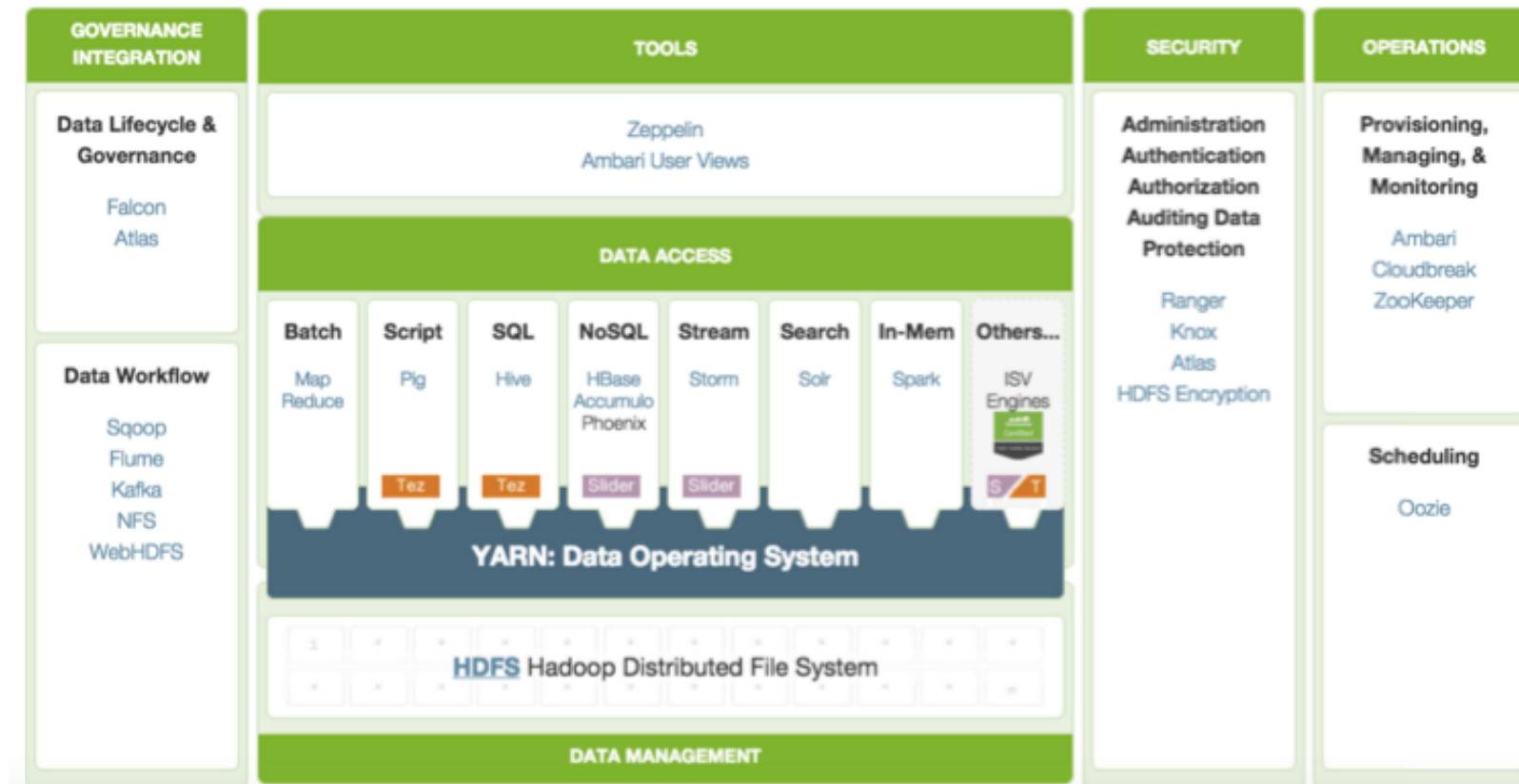
Cloudera



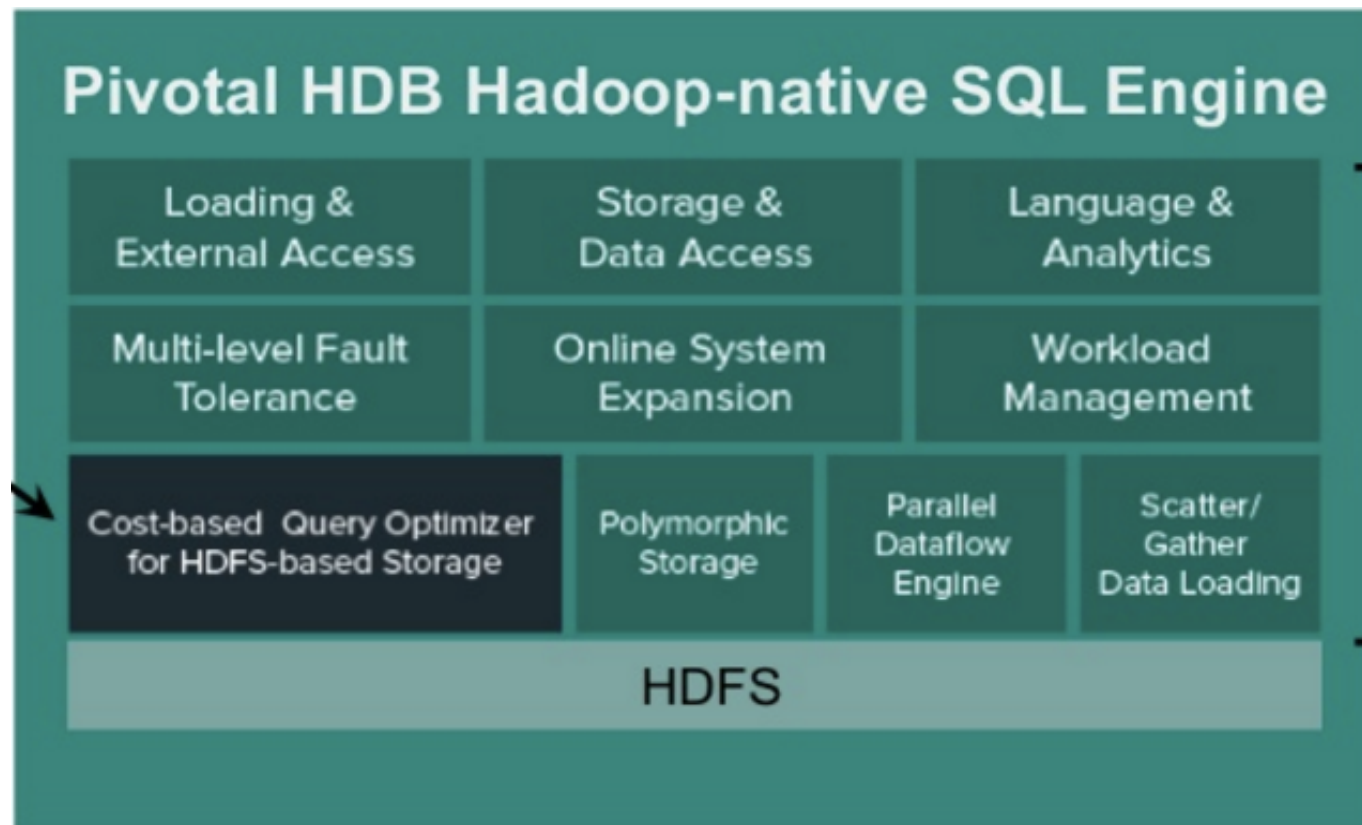
MapR



Hortonworks



Pivotal HDB





IBM InfoSphere BigInsights

Visualization & Discovery

BigSheets

Dashboard &
Visualization

Applications & Development

Apps

Text Analytics

MapReduce

Workflow

Pig & Jaql

Hive

Administration

Admin Console

Monitoring

Integration

JDBC

Netezza

DB2

Streams

DataStage

Guardium

Platform
Computing

Cognos

Flume

Sqoop

Advanced Analytic Engines

Adaptive Algorithms

Text Processing Engine &
Extractor Library)

R

Workload Optimization

Integrated
Installer

Enhanced
Security

Splittable Text
Compression

Adaptive
MapReduce

ZooKeeper

Oozie

Jaql

Flexible
Scheduler

HCatalog

Lucene

Pig

Hive

Index

Runtime / Scheduler

MapReduce

Symphony

Symphony AE

Data Store

HBase

File System

HDFS

GPFS FPO

Management

Security

Audit & History

Lineage

Open Source

IBM

Optional

Amazon EMR

AWS

Services

Edit

IMC Institute

Oregon

Support

Elastic MapReduce

Cluster List

EMR Help

Create cluster

View details

Clone

Terminate

Filter:

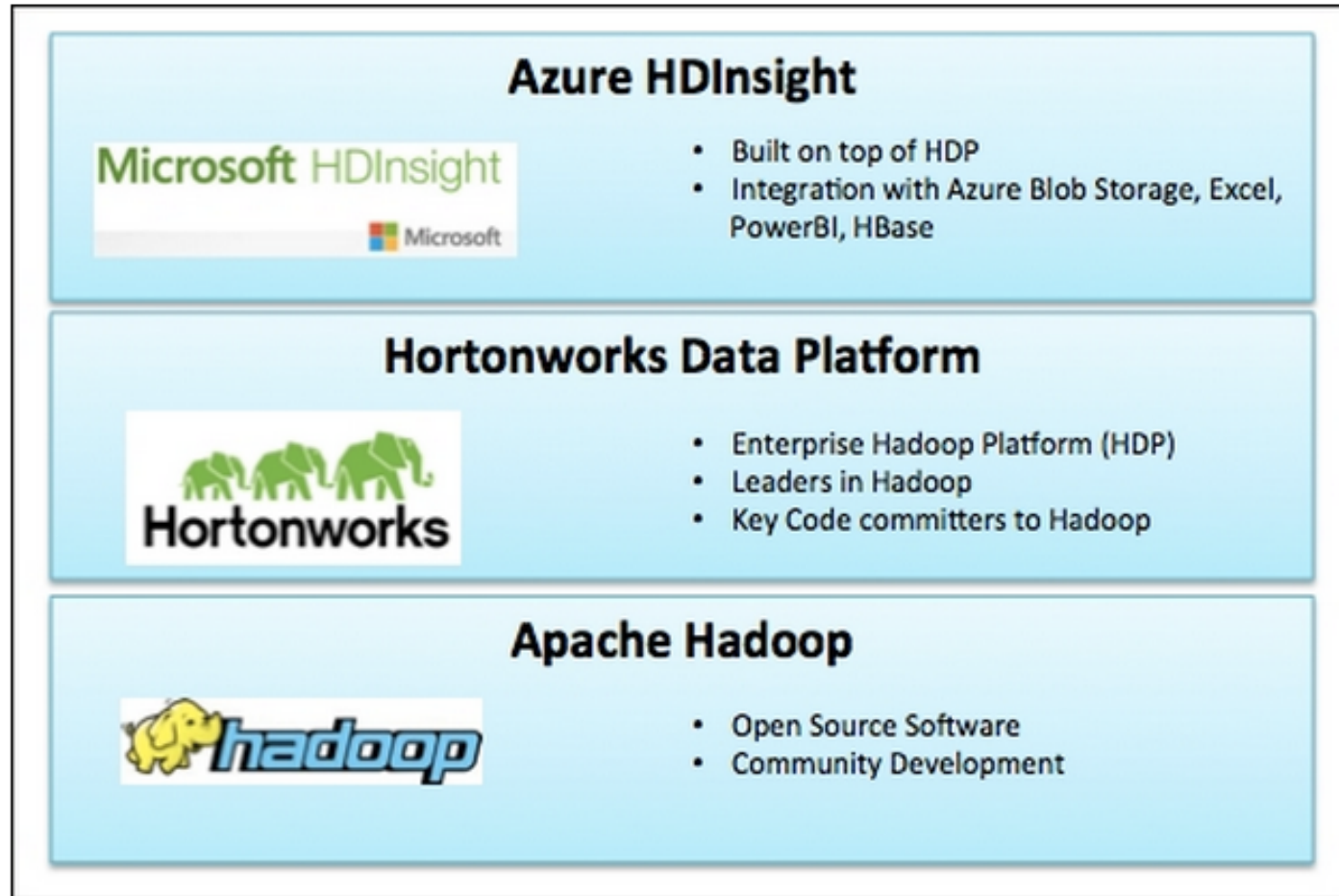
All clusters

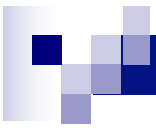
Filter clusters ...

81 clusters (all loaded)

		Name	ID	Status	Creation time (UTC+7)	Elapsed time	Normalized instance ho
				User request			
<input type="checkbox"/>	▶	Sittidate cluster	j-2302SJTZ25UJW	Terminated User request	2016-02-19 09:48 (UTC+7)	5 hours, 24 minutes	48
<input type="checkbox"/>	▶	Plalard cluster	j-36FJTQGN580JF	Terminated User request	2016-02-19 09:48 (UTC+7)	5 hours, 23 minutes	48
<input type="checkbox"/>	▶	Chanoknat cluster	j-3VD0IO2B3E23N	Terminated User request	2016-02-19 09:48 (UTC+7)	5 hours, 23 minutes	48
<input type="checkbox"/>	▶	MDGreat cluster	j-342IJVCW3RNU4	Terminated User request	2016-02-19 09:48 (UTC+7)	5 hours, 23 minutes	48
<input type="checkbox"/>	▶	nutchawit cluster	j-9OI5TYOMONTT	Terminated User request	2016-02-19 09:47 (UTC+7)	5 hours, 24 minutes	48
<input type="checkbox"/>	▶	Wijak cluster	j-21CA76QJLNDQH	Terminated User request	2016-02-19 09:47 (UTC+7)	5 hours, 23 minutes	48

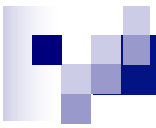
Microsoft Azure HDInsight





Our Works

- Launch a Virtual Server (Ubuntu) on Google Cloud
- Install Docker on Ubuntu
- Pull Cloudera QuickStart to the Docker



Hands-On: Launch a Virtual Server (Ubuntu) on Google Cloud

Cloud.google.com



START YOUR 60 DAY FREE TRIAL

Start building on the same cloud used to power Google, YouTube, and many of the world's largest apps, sites, and services.

- ✓ VIRTUAL MACHINES
- ✓ AUTOSCALING WEB APPLICATIONS
- ✓ DATABASES
- ✓ OBJECT STORAGE
- ✓ NETWORK AND DNS SERVICES
- ✓ BIG DATA
- ✓ ETL AND REALTIME QUERIES
- ✓ TRANSLATION
- ✓ MACHINE LEARNING



Access to all Cloud Platform Products

Get everything you need to build and run your apps, websites and services.



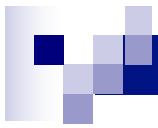
\$300 credit for free

Sign up and get \$300 to spend on Google Cloud Platform over the next 60 days.

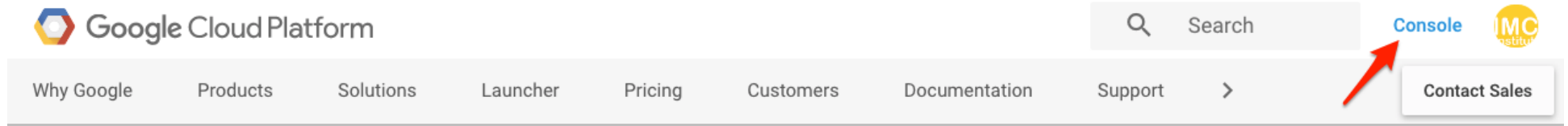


No autocharge after free trial ends

We ask you for your credit card to make sure you are not a robot. You won't be charged during or after your free trial ends.



Click Console



Build What's Next
Better software. Faster.

- ✓ Use Google's core infrastructure, data analytics and machine learning.
- ✓ Secure and fully featured for all enterprises.
- ✓ Committed to open source and industry leading price-performance.

GO TO CONSOLE

CONTACT SALES

Create Project

Google Cloud Platform Project

Getting started

Create project

Try Compute Engine

Create a Linux virtual machine on Google Cloud Compute Engine in this guided walkthrough.

Get started

Use Google APIs

Enable APIs, create credentials, and manage your usage

API Enable and manage APIs

Create an empty project

New Project

Project name ?

Hadoop Project

Your project ID will be hadoop-project-148804 ? Edit

Show advanced options...

CANCEL CREATE

Documentation

- Learn about Compute Engine
- Learn about Cloud Storage
- Learn about App Engine

Name: Hadoop Project

The screenshot shows the Google Cloud Platform (GCP) dashboard interface. At the top, a blue header bar contains the Google Cloud Platform logo, the project name 'Hadoop Project' (which is circled in red), a search bar, and several utility icons. Below the header, a left sidebar lists navigation options: Home, Dashboard (selected), and Activity. The main content area displays four tutorial cards: 'Try Compute Engine' (blue background), 'Learn to use Cloud Storage', 'Try App Engine', and 'Use Google APIs'. Each card includes a brief description and a 'Get started' button.

You have \$300.00 in credit and 55 days left in your free trial. [DISMISS](#) [UPGRADE](#)

Google Cloud Platform **Hadoop Project**

[Home](#) [Dashboard](#) [Activity](#)


Try Compute Engine
Create a Linux virtual machine instance in Compute Engine in this guided walkthrough.
[Get started](#)

Learn to use Cloud Storage
Cloud Storage is a powerful and simple storage service. In this tutorial you'll learn the basics by creating a storage bucket, and then uploading and sharing a sample file as a public URL link.
[Get started](#)

Try App Engine
Create and deploy a Hello World app
[Get started](#)

Use Google APIs
Enable APIs, create credentials, and track your usage
[Get started](#)

Click Compute Engine


 You have \$300.00 in credit and 55 days left in your free trial.

DISMISS

UPGRADE


 Google Cloud Platform


 Filter products & services


 Home

 API API Manager

 Billing

 Cloud Launcher


 Support

 IAM & Admin

COMPUTE

 App Engine

 Compute Engine

 Container Engine

 Networking

board

Try Compute Engine

Create a Linux virtual machine instance in Compute Engine in this guided walkthrough.



Get started

Learn to use Cloud Storage

Cloud Storage is a powerful and simple storage service. In this tutorial you'll learn the basics by creating a storage bucket, and then uploading and sharing a sample file as a public URL link.



Get started

Try App Engine

Create and deploy a Hello World app



Get started

Use Google APIs

Enable APIs, create credentials, and track your usage

API

Enable and manage APIs

Learn Google Cloud Platform

Take an interactive tutorial now and learn how to deploy and build simple applications.

Create a Cloud SQL instance

Cloud SQL is a MySQL database that runs

Create Instance



You have \$300.00 in credit and 55 days left in your free trial.

DISMISS

UPGRADE



Google Cloud Platform Hadoop Project ▾



2



Compute Engine



VM instances



Instance groups



Instance templates



Disks



Snapshots



Images



Metadata



Health checks



Zones



Operations



Quotas



Settings

Compute Engine VM instances

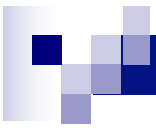
Compute Engine lets you use virtual machines that run on Google's infrastructure. You can choose from micro-VMs to large instances running Debian, Windows, or other standard images. Create your first VM instance or try the quickstart to build a sample app.

Create Instance

or

Take the quickstart





Virtual Server

This lab will use a virtual server to install a Cloudera docker using the following features:

- Name: hadoop-docker
- Zone: asia-east1-c
- Machine type: 4vCPU, 15 GB memory
- Boot disk: Ubuntu **14.04** LTS, **80** GB

Create an Instance



You have \$300.00 in credit and 55 days left in your free trial.

Google Cloud Platform

Hadoop Project

Compute Engine

VM instances

Instance groups

Instance templates

Disks

Snapshots

Images

Metadata

Health checks

Zones

Operations

Quotas

Settings

Create an instance

Name ?

hadoop-docker

Zone ?

asia-east1-c

Machine type

4 vCPUs

15 GB memory

Customize

Upgrade your account to create instances with up to 32 cores

Boot disk ?

New 80 GB standard persistent disk
Image
Ubuntu 14.04 LTS

Change

Identity and API access ?

Service account ?

Compute Engine default service account

Access scopes ?

☒ Allow default access

☐ Allow full access to all Cloud APIs

☐ Set access for each API

Boot disk

Select an image or snapshot to create a boot disk; or attach an existing disk.

OS images

Application images

Custom images

Snapshots

Existing disks

☐ Debian GNU/Linux 8 (jessie)
amd64 built on 2016-10-27

☐ CentOS 6
x86_64 built on 2016-10-27

☐ CentOS 7
x86_64 built on 2016-10-27

☐ CoreOS alpha 1221.0.0
amd64-usr published on 2016-11-03

☐ CoreOS beta 1192.2.0
amd64-usr published on 2016-11-02

☐ CoreOS stable 1185.3.0
amd64-usr published on 2016-11-01

☐ Ubuntu 12.04 LTS
amd64 precise image built on 2016-10-20

☒ Ubuntu 14.04 LTS
amd64 trusty image built on 2016-10-20

☐ Ubuntu 16.04 LTS
amd64 xenial image built on 2016-10-20

☐ Ubuntu 16.10
amd64 yakkety image built on 2016-10-20

☐ Red Hat Enterprise Linux 6
x86_64 built on 2016-10-27

☐ Red Hat Enterprise Linux 7

Boot disk type ?

Standard persistent disk

Size (GB) ?

80

Select

Cancel

Click SSH

Google Cloud Platform Test Hadoop Project

Compute Engine

VM instances

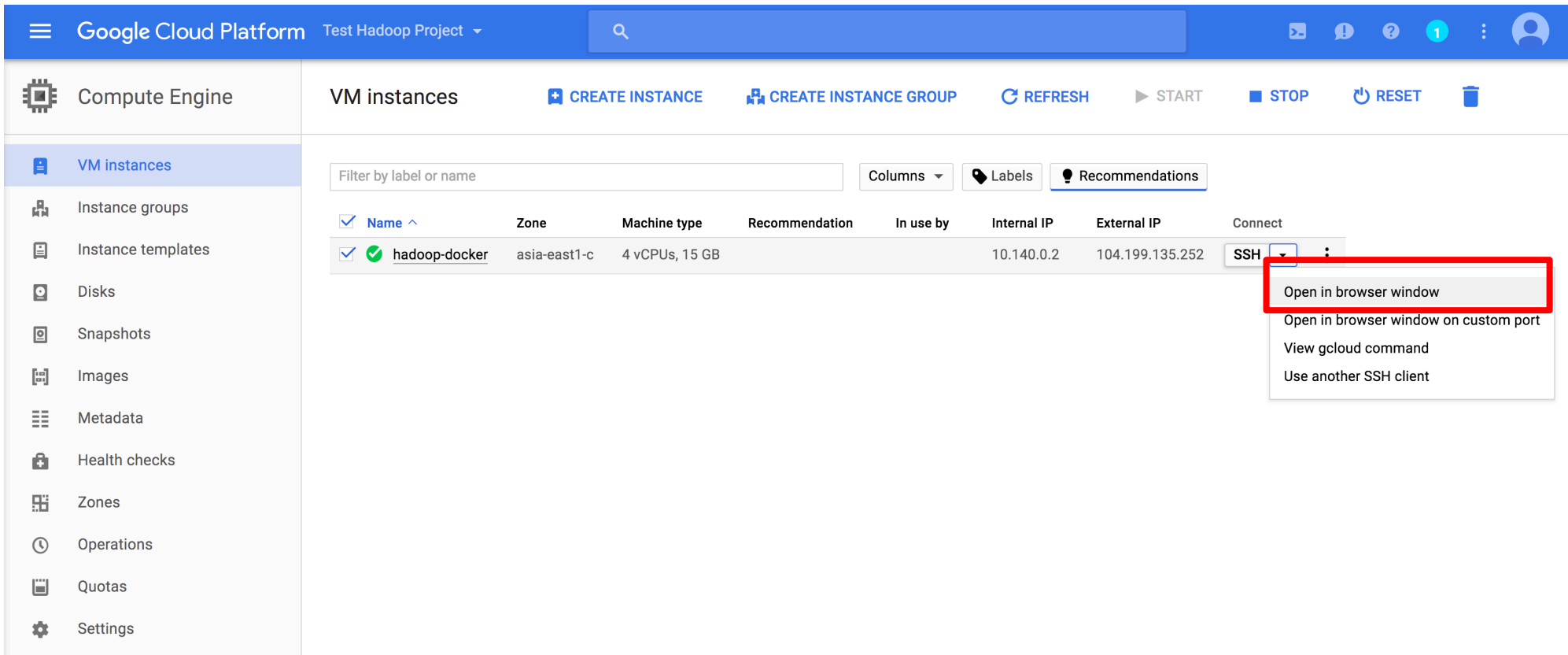
CREATE INSTANCE CREATE INSTANCE GROUP REFRESH START STOP RESET

Filter by label or name Columns Labels Recommendations

Name	Zone	Machine type	Recommendation	In use by	Internal IP	External IP	Connect
hadoop-docker	asia-east1-c	4 vCPUs, 15 GB			10.140.0.2	104.199.135.252	SSH

VM instances sidebar: Instance groups, Instance templates, Disks, Snapshots, Images, Metadata, Health checks, Zones, Operations, Quotas, Settings

Choose Open in browser window



The screenshot shows the Google Cloud Platform interface for the 'Test Hadoop Project'. The left sidebar lists various resources under 'Compute Engine', with 'VM instances' selected. The main panel displays a table of VM instances. The instance 'hadoop-docker' is highlighted, and the 'Connect' dropdown menu is open, showing the option 'Open in browser window' selected and highlighted with a red box.

Google Cloud Platform Test Hadoop Project

Compute Engine

VM instances

CREATE INSTANCE CREATE INSTANCE GROUP REFRESH START STOP RESET

Filter by label or name Columns Labels Recommendations

Name	Zone	Machine type	Recommendation	In use by	Internal IP	External IP	Connect
hadoop-docker	asia-east1-c	4 vCPUs, 15 GB			10.140.0.2	104.199.135.252	SSH Open in browser window Open in browser window on custom port View gcloud command Use another SSH client

Connect to the instance

```
Connected, host fingerprint: ssh-rsa 2048 E1:6D:97:3D:E0:BC:6B:22:56:63:13:7E:CF:41:29:1A
Welcome to Ubuntu 14.04.5 LTS (GNU/Linux 4.4.0-45-generic x86_64)

* Documentation:  https://help.ubuntu.com/

System information as of Mon Nov  7 04:23:21 UTC 2016

System load: 0.08           Memory usage: 0%   Processes:      96
Usage of /:  10.2% of 9.81GB Swap usage:   0%   Users logged in: 0

Graph this data and manage this system at:
https://landscape.canonical.com/

Get cloud support with Ubuntu Advantage Cloud Guest:
http://www.ubuntu.com/business/services/cloud

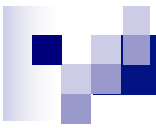
0 packages can be updated.
0 updates are security updates.

Your Hardware Enablement Stack (HWE) is supported until April 2019.

The programs included with the Ubuntu system are free software;
the exact distribution terms for each program are described in the
individual files in /usr/share/doc/*/copyright.

Ubuntu comes with ABSOLUTELY NO WARRANTY, to the extent permitted by
applicable law.

contact@hadoop-docker:~$
```

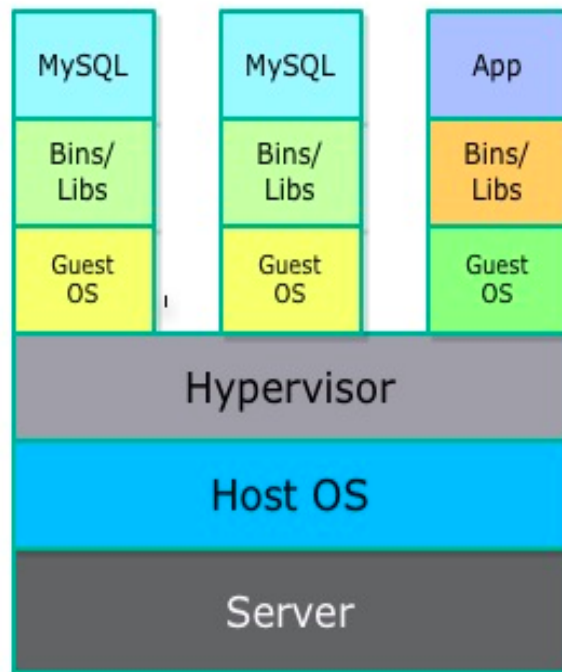


Hands-On: Install a Docker Engine

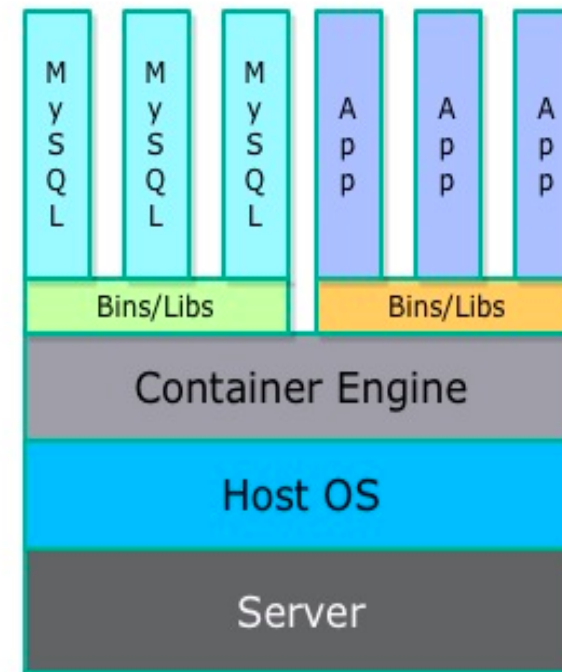


Docker v.s. Hypervisor

Virtual Machines



Containers





Update OS (Ubuntu)

- Command: `sudo apt-get update`

```
contact@hadoop-docker:~$ sudo apt-get update
Ign http://asia-east1.gce.archive.ubuntu.com trusty InRelease
Get:1 http://asia-east1.gce.archive.ubuntu.com trusty-updates InRelease [65.9 kB]
Get:2 http://asia-east1.gce.archive.ubuntu.com trusty-backports InRelease [65.9 kB]
Hit http://asia-east1.gce.archive.ubuntu.com trusty Release.gpg
Hit http://asia-east1.gce.archive.ubuntu.com trusty Release
Get:3 http://asia-east1.gce.archive.ubuntu.com trusty-updates/main Sources [384 kB]
Get:4 http://asia-east1.gce.archive.ubuntu.com trusty-updates/restricted Sources [5,888 B]
Get:5 http://asia-east1.gce.archive.ubuntu.com trusty-updates/universe Sources [169 kB]
Get:6 http://asia-east1.gce.archive.ubuntu.com trusty-updates/multiverse Sources [7,531 B]
Get:7 http://asia-east1.gce.archive.ubuntu.com trusty-updates/main amd64 Packages [913 kB]
Get:8 http://asia-east1.gce.archive.ubuntu.com trusty-updates/restricted amd64 Packages [16.4 kB]
Get:9 http://asia-east1.gce.archive.ubuntu.com trusty-updates/universe amd64 Packages [388 kB]
Get:10 http://asia-east1.gce.archive.ubuntu.com trusty-updates/multiverse amd64 Packages [14.0 kB]
Get:11 http://asia-east1.gce.archive.ubuntu.com trusty-updates/main Translation-en [443 kB]
```

Docker Installation

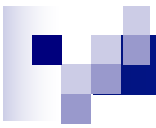
- Command: `sudo apt-get install docker.io`

```
contact@hadoop-docker:~$ sudo apt-get install docker.io
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following extra packages will be installed:
  aufs-tools cgroup-lite git git-man liberror-perl
Suggested packages:
  btrfs-tools debootstrap lxc rinse git-daemon-run git-daemon-sysvinit git-doc
  git-el git-email git-gui gitk gitweb git-arch git-bzr git-cvs git-mediawiki
  git-svn
The following NEW packages will be installed:
  aufs-tools cgroup-lite docker.io git git-man liberror-perl
0 upgraded, 6 newly installed, 0 to remove and 17 not upgraded.
Need to get 8,150 kB of archives.
After this operation, 51.4 MB of additional disk space will be used.
Do you want to continue? [Y/n] Y
Get:1 http://asia-east1.gce.archive.ubuntu.com/ubuntu/ trusty/universe aufs-tools amd64 1:3
Get:2 http://asia-east1.gce.archive.ubuntu.com/ubuntu/ trusty-updates/universe docker.io am
Get:3 http://asia-east1.gce.archive.ubuntu.com/ubuntu/ trusty/main liberror-perl all 0.17-1
```



Docker commands:

- docker images
- docker ps
- docker attach id
- docker kill id
- docker commit id
- Exit from container
 - exit (exit & kill the running image)
 - Ctrl-P, Ctrl-Q (exit without killing the running image)



Install Cloudera Quickstart on Docker Container



Pull Cloudera Quickstart

\$ sudo docker pull cloudera/quickstart:latest

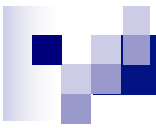
```
contact@hadoop-docker:~$ sudo docker pull cloudera/quickstart:latest
2cda82941cb7: Already exists
Digest: sha256:f91bee4cdfa2c92ea3652929a22f729d4d13fc838b00f120e630f91c941acb63 1.157 GB/4.444 GB
Status: Downloaded newer image for cloudera/quickstart:latest
```



Verify the image was successfully pulled

\$ sudo docker images

```
contact@hadoop-docker:~$ sudo docker images
REPOSITORY          TAG                 IMAGE ID            CREATED             VIRTUAL SIZE
cloudera/quickstart  latest             2cda82941cb7       7 months ago       6.336 GB
contact@hadoop-docker:~$
```



Running Cloudera Docker On Google Cloud

Run Cloudera quickstart

```
$ sudo docker run --hostname=quickstart.cloudera  
--privileged=true -t -i [OPTIONS] [IMAGE] /usr/bin/docker-quickstart
```

Example: `sudo docker run --hostname=quickstart.cloudera
--privileged=true -t -i -p 8888:8888 cloudera/quickstart
/usr/bin/docker-quickstart`

```
contact@hadoop-docker:~$ sudo docker run --hostname=quickstart.cloudera --privileged=true -t -i -p 8888:8888 cloudera/quickstart /usr/bin/docker-quickstart  
Starting mysqld: [ OK ]  
  
if [ "$1" == "start" ] ; then  
    if [ "${EC2}" == 'true' ]; then  
        FIRST_BOOT_FLAG=/var/lib/cloudera-quickstart/.ec2-key-installed  
        if [ ! -f "${FIRST_BOOT_FLAG}" ]; then  
            METADATA_API=http://169.254.169.254/latest/meta-data  
            KEY_URL=${METADATA_API}/public-keys/0/openssh-key  
            SSH_DIR=/home/cloudera/.ssh  
            mkdir -p ${SSH_DIR}  
            chown cloudera:cloudera ${SSH_DIR}  
            curl ${KEY_URL} >> ${SSH_DIR}/authorized_keys  
            touch ${FIRST_BOOT_FLAG}
```


Successful running the Cloudera image

```
Using CATALINA_BASE:  /var/lib/oozie/tomcat-deployment
Using CATALINA_HOME:  /usr/lib/bigtop-tomcat
Using CATALINA_TMPDIR: /var/lib/oozie
Using JRE_HOME:       /usr/java/jdk1.7.0_67-cloudera
Using CLASSPATH:      /usr/lib/bigtop-tomcat/bin/bootstrap.jar
Using CATALINA_PID:   /var/run/oozie/oozie.pid
Starting Solr server daemon: [ OK ]
Using CATALINA_BASE:  /var/lib/solr/tomcat-deployment
Using CATALINA_HOME:  /usr/lib/solr/../bigtop-tomcat
Using CATALINA_TMPDIR: /var/lib/solr/
Using JRE_HOME:       /usr/java/jdk1.7.0_67-cloudera
Using CLASSPATH:      /usr/lib/solr/../bigtop-tomcat/bin/bootstrap.jar
Using CATALINA_PID:   /var/run/solr/solr.pid
Started Impala Catalog Server (catalogd) : [ OK ]
Started Impala Server (impalad): [ OK ]
[root@quickstart /]#
```

Select the instance

Google Cloud Platform Hadoop Project

Compute Engine

VM instances

CREATE INSTANCE CREATE INSTANCE GROUP REFRESH START STOP

Filter by label or name Columns Labels Recommendations

CPU utilization

1 hour 6 hours 12 hours 1 day 2 days 4 days 7 days 14 days 30 days


CPU

% CPU

40 30 20 10

7. Nov, 11:00 AM 7. Nov, 11:15 AM 7. Nov, 11:30 AM 7. Nov, 11:45 AM 7. Nov, 11:59 AM

CPU: 2.446

<input type="checkbox"/> Name ^	Zone	Machine type	Recommendation	In use by	Internal IP	External IP	Connect
<input checked="" type="checkbox"/>  hadoop-docker	asia-east1-c	4 vCPUs, 15 GB			10.140.0.2	104.155.230.62	SSH

Click Network default

Google Cloud Platform Hadoop Project

Compute Engine

VM instances

Instance groups

Instance templates

Disks

Snapshots

Images

Metadata

Health checks

Zones

Operations

Quotas

Settings

n1-standard-4 (4 vCPUs, 15 GB memory)

CPU platform
Intel Ivy Bridge

Zone
asia-east1-c

External IP
104.155.230.62 (ephemeral)

Internal IP
10.140.0.2

IP forwarding
off

Boot disk and local disks

Name	Size (GB)	Type	Mode
hadoop-docker	80	Standard persistent disk	Boot, read/write

☒ Delete boot disk when instance is deleted

Additional disks
None

Network
[default](#)

Subnetwork
default

Click Add firewall rule

Google Cloud Platform Hadoop Project

Networking

Networks

External IP addresses

Firewall rules

Routes

Load balancing

Cloud DNS

VPN

Cloud Routers

Network details

EDIT

DELETE NETWORK

Name ^	Region	IP address ranges	Gateway
default	us-central1	10.128.0.0/20	10.128.0.1
default	europa-west1	10.132.0.0/20	10.132.0.1
default	us-west1	10.138.0.0/20	10.138.0.1
default	asia-east1	10.140.0.0/20	10.140.0.1
default	us-east1	10.142.0.0/20	10.142.0.1
default	asia-northeast1	10.146.0.0/20	10.146.0.1

Firewall rule

Add firewall rule

Delete

Name ^	Source tag / IP range / Subnetworks	Allowed protocols / ports	Target tags
<input type="checkbox"/> default-allow-icmp	0.0.0.0/0	icmp	Apply to all targets
<input type="checkbox"/> default-allow-internal	10.128.0.0/9	tcp:0-65535, 2 more	Apply to all targets
<input type="checkbox"/> default-allow-rdp	0.0.0.0/0	tcp:3389	Apply to all targets
<input type="checkbox"/> default-allow-ssh	0.0.0.0/0	tcp:22	Apply to all targets



A Firewall Rule

- Name: **hue**
- Source IP ranges: **0.0.0.0/0**
- Allowed protocols and ports: **tcp:8888**

Create a firewall rule

Google Cloud Platform Hadoop Project

Networking

- Networks
- External IP addresses
- Firewall rules**
- Routes
- Load balancing
- Cloud DNS
- VPN
- Cloud Routers

Create a firewall rule

By default, incoming traffic from outside your network is blocked. To allow incoming traffic, set up a firewall rule. Firewall rules regulate only incoming traffic to an instance. When a connection is established with an instance, traffic is permitted in both directions over that connection. [Learn more](#)

Name ?

hue

Description (Optional)

Network ?

default

Source filter ?

IP ranges

Source IP ranges ?

0.0.0.0/0 x

Allowed protocols and ports ?

tcp:8888

Target tags (Optional) ?

Create Cancel

See New Firewall Rule

Google Cloud Platform Hadoop Project

Networking

- Networks
- External IP addresses
- Firewall rules
- Routes
- Load balancing
- Cloud DNS
- VPN
- Cloud Routers

Network details

EDIT DELETE NETWORK

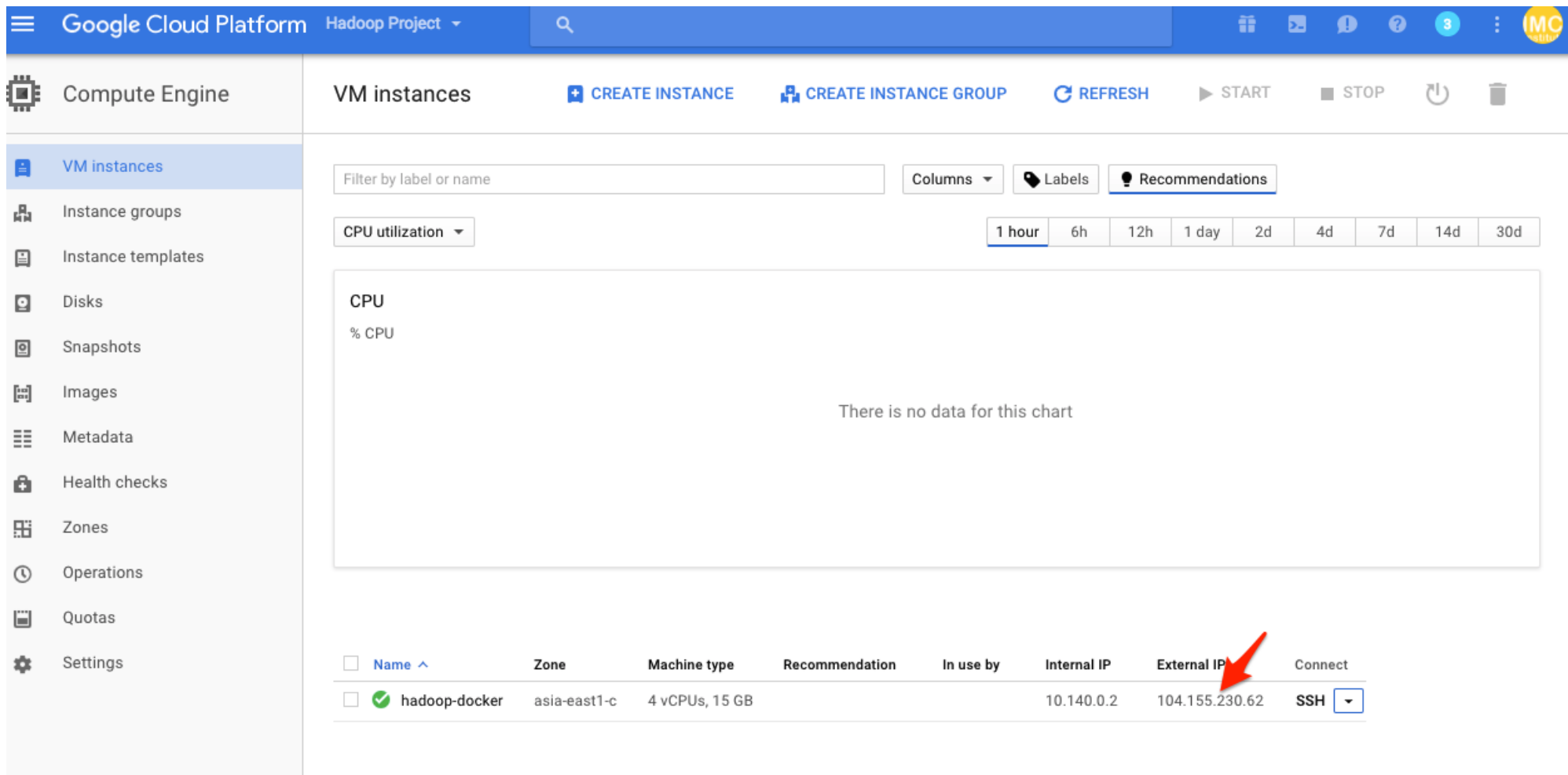
default	europa-west1	10.132.0.0/20	10.132.0.1
default	us-west1	10.138.0.0/20	10.138.0.1
default	asia-east1	10.140.0.0/20	10.140.0.1
default	us-east1	10.142.0.0/20	10.142.0.1
default	asia-northeast1	10.146.0.0/20	10.146.0.1

Firewall rules

Add firewall rule Delete

Name	Source tag / IP range / Subnetworks	Allowed protocols / ports	Target tags
<input type="checkbox"/> default-allow-icmp	0.0.0.0/0	icmp	Apply to all targets
<input type="checkbox"/> default-allow-internal	10.128.0.0/9	tcp:0-65535, 2 more	Apply to all targets
<input type="checkbox"/> default-allow-rdp	0.0.0.0/0	tcp:3389	Apply to all targets
<input type="checkbox"/> default-allow-ssh	0.0.0.0/0	tcp:22	Apply to all targets
<input type="checkbox"/> hue	0.0.0.0/0	tcp:8888	Apply to all targets

Finding the instance's external IP address

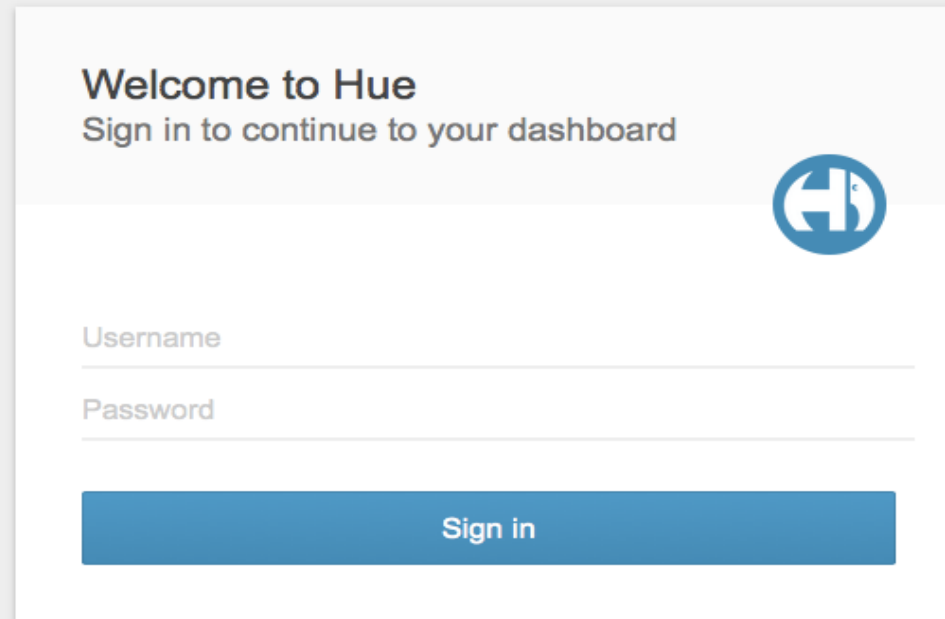


The screenshot shows the Google Cloud Platform interface for VM instances. The left sidebar contains the 'Compute Engine' menu with options like 'VM instances', 'Instance groups', 'Instance templates', 'Disks', 'Snapshots', 'Images', 'Metadata', 'Health checks', 'Zones', 'Operations', 'Quotas', and 'Settings'. The main panel is titled 'VM instances' and includes buttons for 'CREATE INSTANCE', 'CREATE INSTANCE GROUP', 'REFRESH', 'START', 'STOP', and a trash icon. Below these are filters for 'Filter by label or name', 'Columns', 'Labels', and 'Recommendations'. A 'CPU utilization' chart is shown with a message 'There is no data for this chart'. At the bottom, a table lists VM instances with columns: Name, Zone, Machine type, Recommendation, In use by, Internal IP, External IP, and Connect. A red arrow points to the 'External IP' column for the instance 'hadoop-docker'.

Name	Zone	Machine type	Recommendation	In use by	Internal IP	External IP	Connect
hadoop-docker	asia-east1-c	4 vCPUs, 15 GB			10.140.0.2	104.155.230.62	SSH

Login to Hue: <http://external-ip-address:8888>

Username: cloudera Password: cloudera

The image shows the Hue login interface. It features a white login box centered on a light gray background. Inside the box, the text "Welcome to Hue" is displayed in a bold, dark font, followed by "Sign in to continue to your dashboard" in a smaller, lighter font. To the right of this text is the Hue logo, which is a blue circle containing a white stylized 'H' and a small bird icon. Below the text and logo are two input fields: "Username" and "Password", each with a horizontal line for text entry. At the bottom of the box is a large blue button with the text "Sign in" in white.

Welcome to Hue

Sign in to continue to your dashboard

Username

Password

Sign in

Hue and the Hue logo are trademarks of Cloudera, Inc.



Quick Start Wizard - Hue™ 3.9.0 - The Hadoop UI

Step 1: Check Configuration Step 2: Examples Step 3: Users Step 4: Goals

Checking current configuration

Configuration files located in `/etc/hue/conf.empty`

All OK. Configuration check passed.

[Back](#) [Next](#)

Hue and the Hue logo are trademarks of Cloudera, Inc.