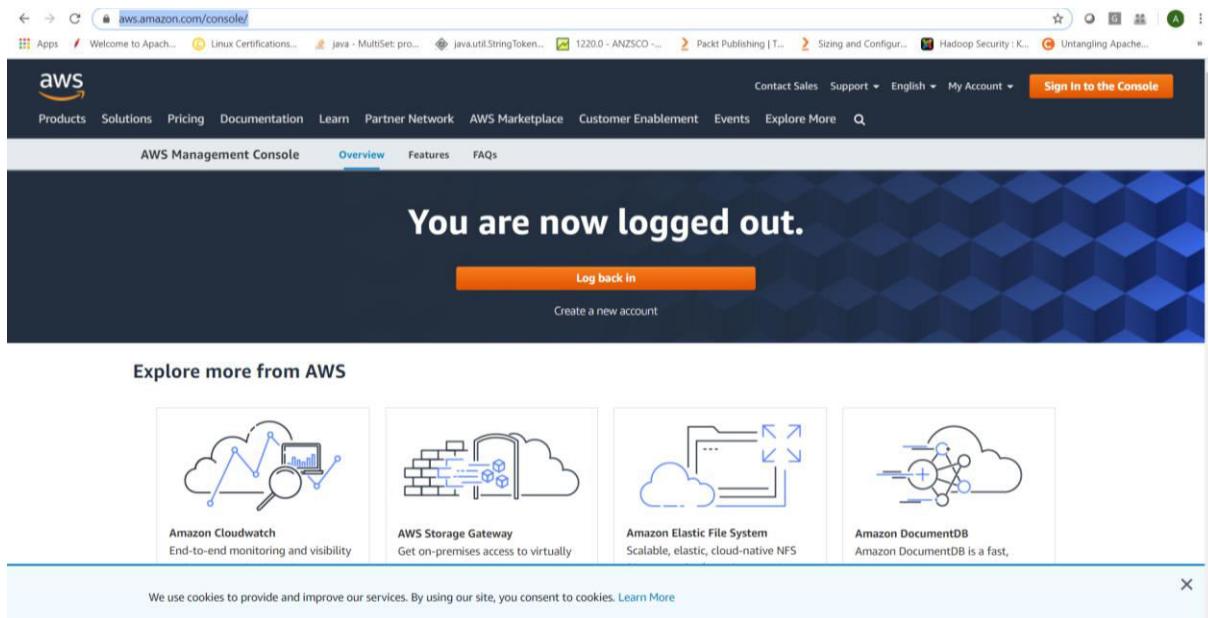


Setup Cloudera Cluster on AWS

Create an account for aws services on <https://aws.amazon.com/console/>



& sign into console



AWS Management Console

console.aws.amazon.com/console/home?nc2=h_ct&src=header-signin®ion=us-east-1

AWS Services

Servies Resource Groups

AWS Management Console

AWS services

Find Services You can enter names, keywords or acronyms. Example: Relational Database Service, database, RDS

Recently visited services VPC AWS Glue S3
EC2 EMR

All services

- Compute
 - EC2
 - Lightsail
 - Lambda
 - Batch
 - Elastic Beanstalk
 - Serverless Application Repository
 - AWS Outposts
 - EC2 Image Builder
- Blockchain
 - Amazon Managed Blockchain
- Satellite
 - Ground Station
- Quantum Technologies
 - Amazon Braket
- Security, Identity, & Compliance
 - IAM
 - Resource Access Manager
 - Cognito
 - Secrets Manager
 - GuardDuty
 - Inspector
 - Amazon Macie

Stay connected to your AWS resources on-the-go

Download the AWS Console Mobile App to your iOS or Android mobile device. Learn more

Explore AWS

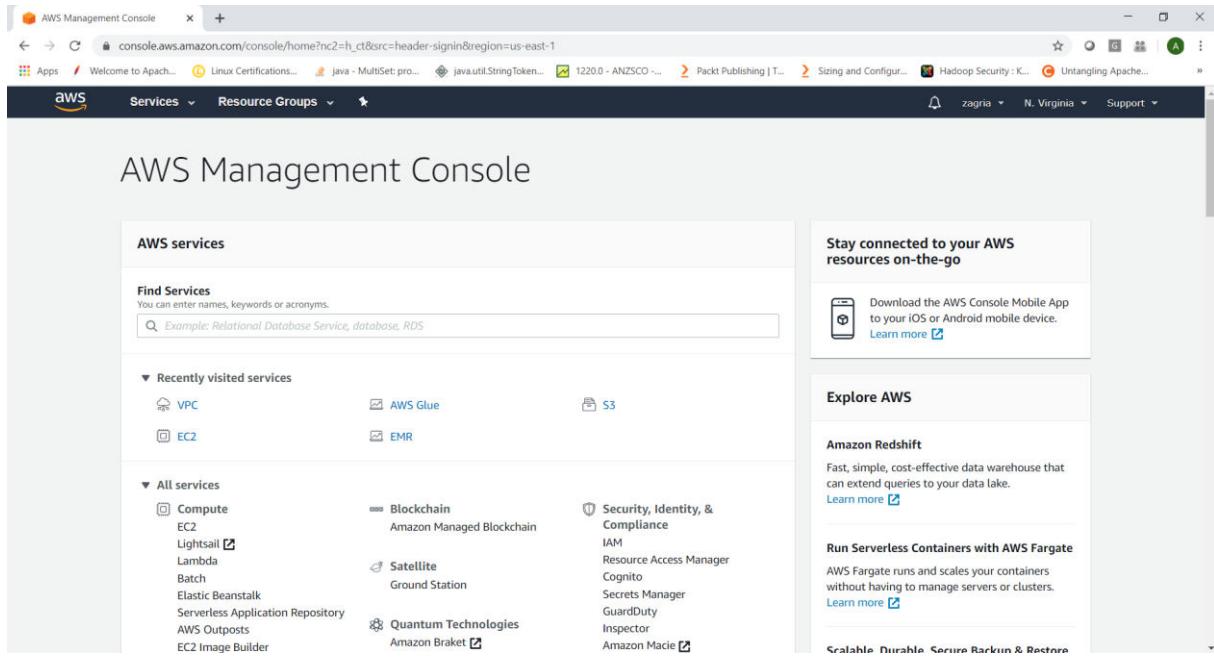
Amazon Redshift

Fast, simple, cost-effective data warehouse that can extend queries to your data lake. Learn more

Run Serverless Containers with AWS Fargate

AWS Fargate runs and scales your containers without having to manage servers or clusters. Learn more

Scalable Durable Secure Backup & Restore



Create Default VPC (if not existing)

EC2 Management Console

console.aws.amazon.com/ec2/v2/home?region=us-east-1#Home:

Services Resource Groups

History EC2 Console Home VPC AWS Glue EMR S3

vpc

VPC Isolated Cloud Resources

AWS Firewall Manager Central management of firewall rules

Detective Investigate and analyze potential security issues

Route 53 Resolver Route 53 Resolver provides recursive DNS for your Amazon VPC and on-premises networks over VPN or Direct Connect.

AWS Outposts EC2 Image Builder Quantum Technologies Amazon Braket Data Pipeline AWS Data Exchange AWS Glue AWS Lake Formation MSK

Storage S3 EFS FSx S3 Glacier Storage Gateway AWS Backup

Management & Governance AWS Organizations CloudWatch AWS Auto Scaling CloudFormation CloudTrail Config OpsWorks Service Catalog

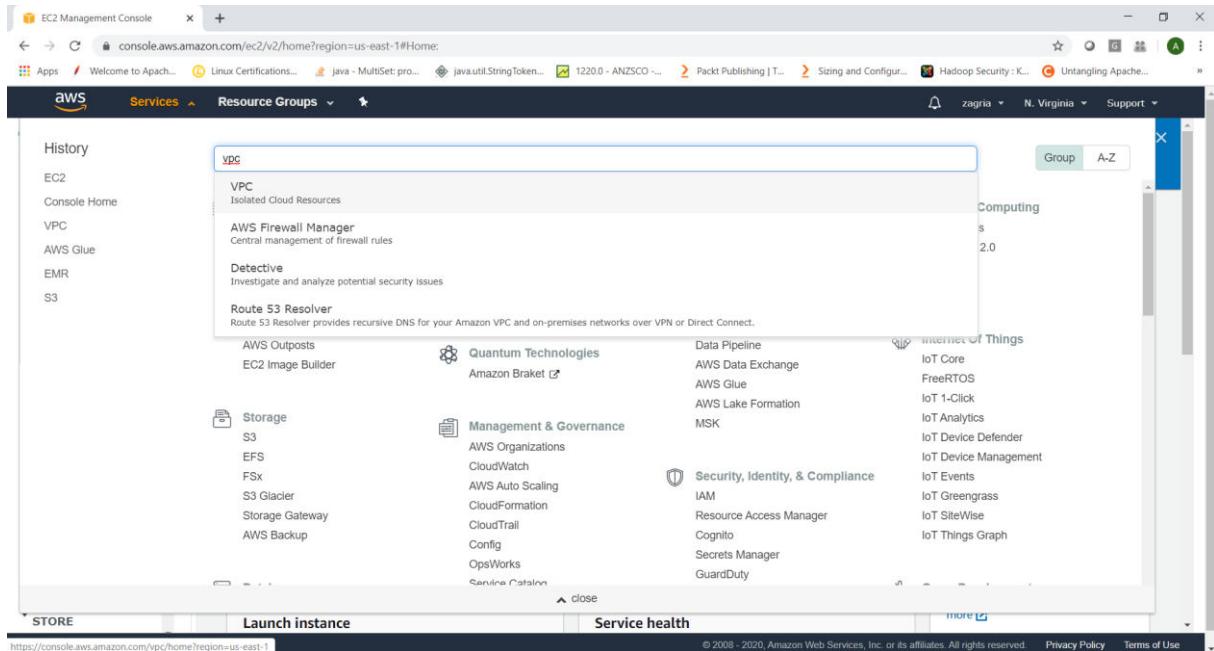
Internet of Things IoT Core FreeRTOS IoT 1-Click IoT Analytics IoT Device Defender IoT Device Management IoT Events IoT Greengrass IoT SiteWise IoT Things Graph

close

STORE Launch instance Service health more

https://console.aws.amazon.com/vpc/home?region=us-east-1

© 2006–2020, Amazon Web Services, Inc. or its affiliates. All rights reserved. Privacy Policy Terms of Use



VPC Management Console

Welcome to Apache... Linux Certifications... Java - MultiSet: pro... java.util.StringToken... 1220.0 - ANZSCO ... Packt Publishing | T... Sizing and Configur... Hadoop Security : K... Untangling Apache...

AWS Services Resource Groups

VPC Dashboard

Filter by VPC: Select a VPC

VIRTUAL PRIVATE CLOUD

- Your VPCs
- Subnets
- Route Tables
- Internet Gateways
- Egress Only Internet Gateways
- DHCP Options Sets
- Elastic IPs
- Endpoints
- Endpoint Services
- NAT Gateways
- Peering Connections

SECURITY

- Network ACLs
- Security Groups

VIRTUAL PRIVATE NETWORK (VPN)

- Customer Gateways
- Virtual Private Gateways

Resources by Region Refresh Resources

Note: Your Instances will launch in the US East (N. Virginia) region.

VPCs See all regions	N. Virginia 0	NAT Gateways See all regions	N. Virginia 0
Subnets See all regions	N. Virginia 0	VPC Peering Connections See all regions	N. Virginia 0
Route Tables See all regions	N. Virginia 0	Network ACLs See all regions	N. Virginia 0
Internet Gateways See all regions	N. Virginia 0	Security Groups See all regions	N. Virginia 0
Egress-only Internet Gateways See all regions	N. Virginia 0	Customer Gateways See all regions	N. Virginia 0
DHCP options sets See all regions	N. Virginia 1	Virtual Private Gateways See all regions	N. Virginia 0

Service Health

Current Status	Details
Green Amazon EC2 - US East (N. Virginia)	Service is operating normally

View complete service health details

Account Attributes

Resource ID length management

Additional Information

VPC Documentation All VPC Resources Forums Report an Issue

Transit Gateway Network Manager

Network Manager enables centrally manage your global network across AWS and on-premises. Learn more

Get started with Network Manager

Feedback English (US)

© 2008 - 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved. Privacy Policy Terms of Use

Click on VPCs

vpcs | VPC Management Console

Welcome to Apache... Linux Certifications... Java - MultiSet: pro... java.util.StringToken... 1220.0 - ANZSCO ... Packt Publishing | T... Sizing and Configur... Hadoop Security : K... Untangling Apache...

AWS Services Resource Groups

VPC Dashboard

Filter by VPC: Select a VPC

YOUR VPCs

- Subnets
- Route Tables
- Internet Gateways
- Egress Only Internet Gateways
- DHCP Options Sets
- Elastic IPs
- Endpoints
- Endpoint Services
- NAT Gateways
- Peering Connections

SECURITY

- Network ACLs
- Security Groups

VIRTUAL PRIVATE NETWORK (VPN)

- Customer Gateways
- Virtual Private Gateways

Create VPC Actions

Actions ▾

- Create Default VPC
- Create flow log
- Edit DHCP options set
- Edit DNS resolution
- Edit DNS hostnames
- Add/Edit Tags

Filter by tag keyword

You do not have any VPCs in this region

Click the Create VPC button to create your first VPC

Create VPC

Feedback English (US)

© 2008 - 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved. Privacy Policy Terms of Use

The screenshot shows the AWS VPC Management Console with the 'Create Default VPC' wizard. The steps completed are:

- Step 1: Set VPC Name (Completed)
- Step 2: Set Subnet Range (Completed)
- Step 3: Set Route Table (Completed)
- Step 4: Set Internet Gateway (Completed)
- Step 5: Review and Create (In Progress)

A success message box is displayed at the bottom left:

✓ The following Default VPC was created:

VPC ID: vpc-0fc65fb524e2580d1

Create

Now Create a Key Pair which will be used to access your instances .(Once created i.e. xyz.pem download it)

The screenshot shows the AWS Management Console with the 'vpcs | VPC Management Console' tab selected. The Services menu is open, and the 'Compute' section is expanded, showing the following services:

- EC2
- Lightsail
- Lambda
- Batch
- Elastic Beanstalk
- Serverless Application Repository
- AWS Outposts
- EC2 Image Builder

EC2 Management Console

Welcome to the new EC2 console! We're redesigning the EC2 console to make it easier to use and improve performance. We'll release new screens periodically. We encourage you to try them and let us know where we can make improvements. To switch between the old console and the new console, use the New EC2 Experience toggle.

Resources

You are using the following Amazon EC2 resources in the US East (N. Virginia) Region:

Running instances	0	Elastic IPs	0
Dedicated Hosts	0	Snapshots	0
Volumes	0	Load balancers	0
Key pairs	1	Security groups	1
Placement groups	0		

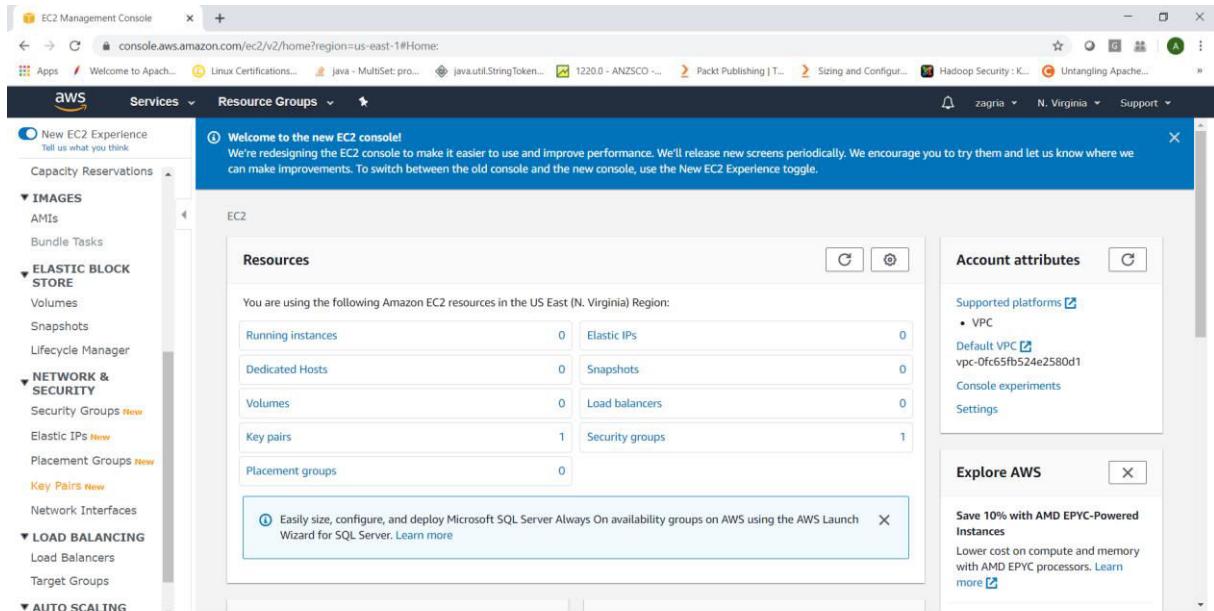
Easily size, configure, and deploy Microsoft SQL Server Always On availability groups on AWS using the AWS Launch Wizard for SQL Server. Learn more

Account attributes

Supported platforms: VPC
Default VPC: vpc-0fc65fb524e2580d1
Console experiments
Settings

Explore AWS

Save 10% with AMD EPYC-Powered Instances
Lower cost on compute and memory with AMD EPYC processors. Learn more



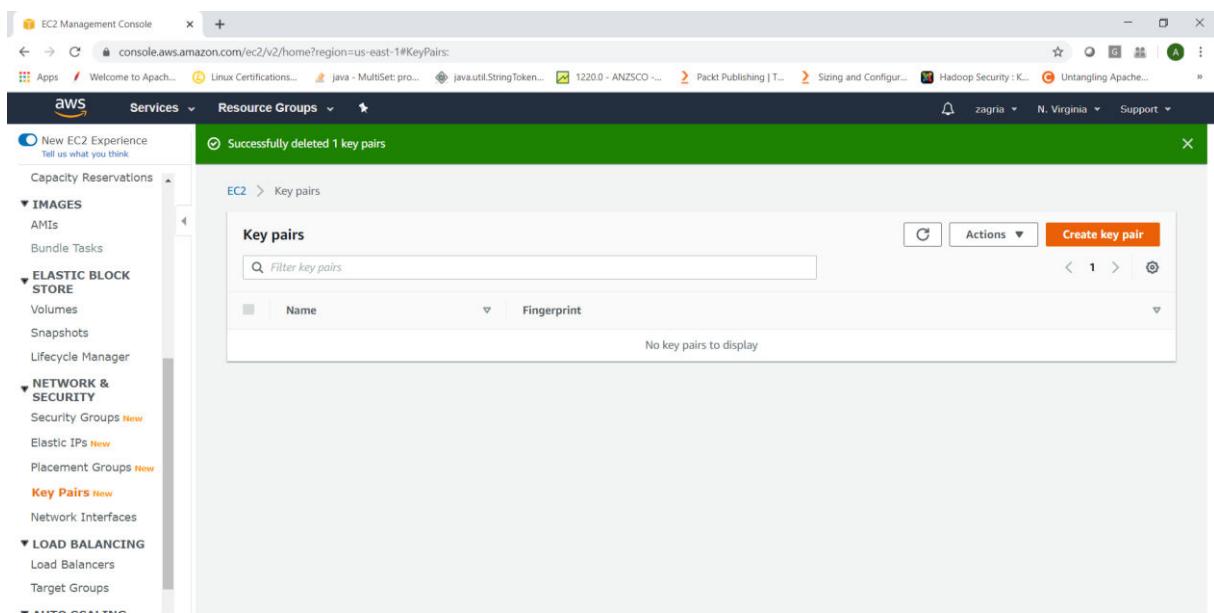
EC2 Management Console

Successfully deleted 1 key pairs

Key pairs

Filter key pairs

Name	Fingerprint
No key pairs to display	



Create key pair | EC2 Management Console

EC2 > Key pairs > Create key pair

Create key pair

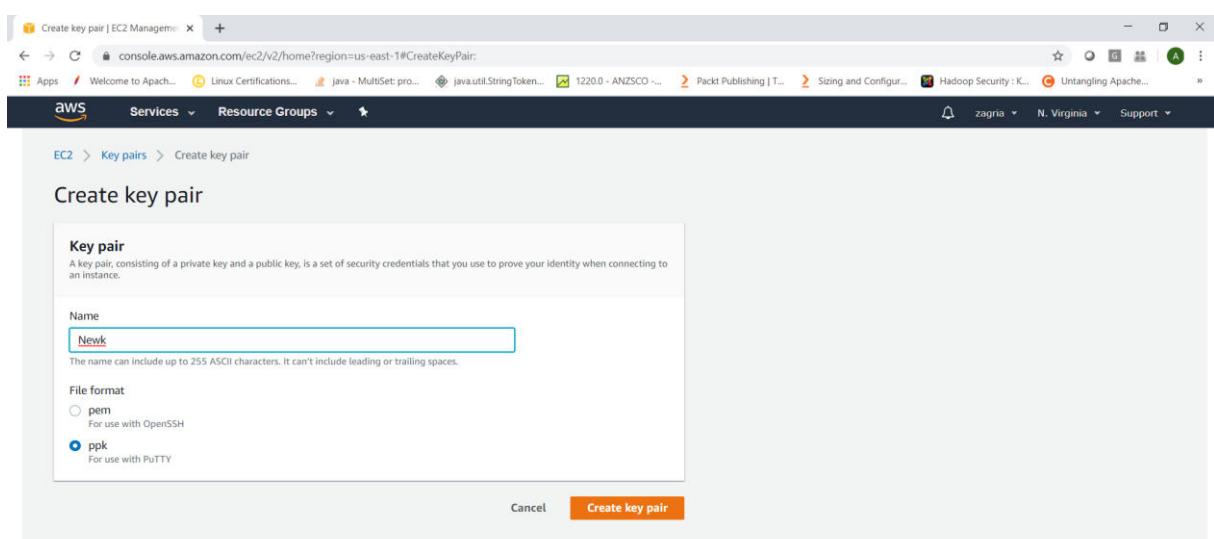
Key pair
A key pair, consisting of a private key and a public key, is a set of security credentials that you use to prove your identity when connecting to an instance.

Name

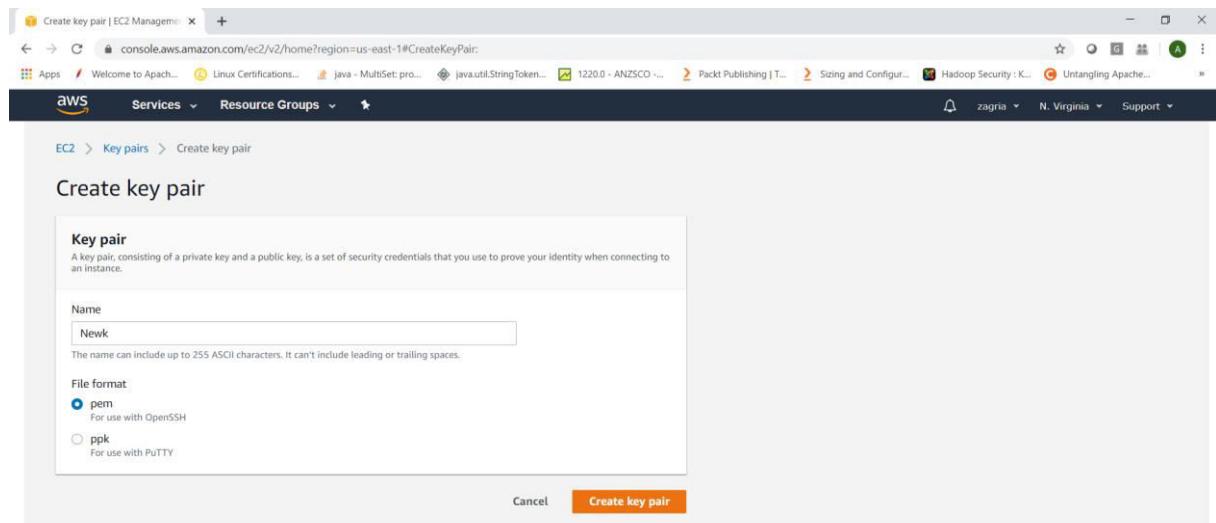
The name can include up to 255 ASCII characters. It can't include leading or trailing spaces.

File format
 pem
For use with OpenSSH
 ppk
For use with PuTTY

Create key pair



Or



The screenshot shows the 'Create key pair' dialog box. It includes a 'Key pair' description, a 'Name' input field containing 'Newk', a 'File format' section with 'pem' selected, and a 'Create key pair' button.

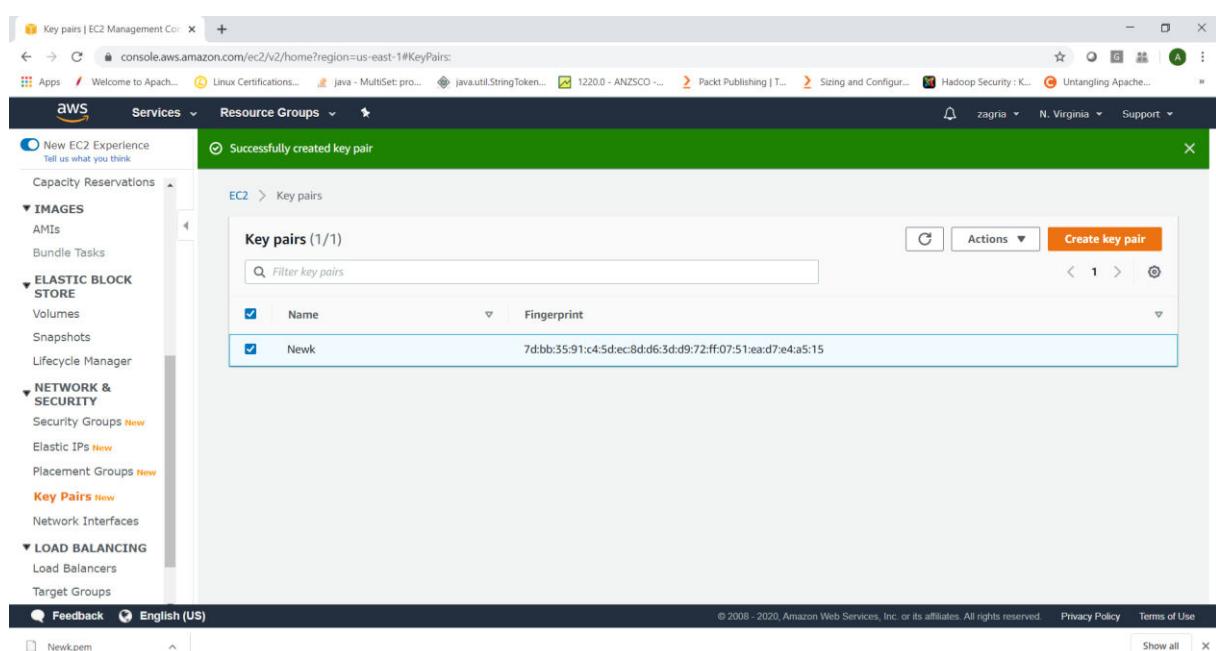
Key pair
A key pair, consisting of a private key and a public key, is a set of security credentials that you use to prove your identity when connecting to an instance.

Name

The name can include up to 255 ASCII characters. It can't include leading or trailing spaces.

File format
 pem
For use with OpenSSH
 ppk
For use with PuTTY

Cancel **Create key pair**



The screenshot shows the 'Key pairs' list page. A green success message at the top says 'Successfully created key pair'. The table lists one key pair named 'Newk' with the fingerprint '7dbb:35:91:c4:5d:ec:8d:d6:3d:d9:72:ff:07:51:ea:d7:e4:a5:15'.

Name	Fingerprint
Newk	7dbb:35:91:c4:5d:ec:8d:d6:3d:d9:72:ff:07:51:ea:d7:e4:a5:15

Feedback English (US) © 2008 - 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved. Privacy Policy Terms of Use Show all

Now Create a Security Group which will allow incoming and outgoing traffic for your instances.

The screenshot shows the AWS EC2 Management Console. The left sidebar has 'Key pairs' selected under 'EC2'. A green banner at the top right says 'Successfully created key pair'. The main area shows a table titled 'Key pairs (1/1)' with one row for 'Newk'. The row includes a checkbox, the name 'Newk', and a 'Fingerprint' column showing '7dbb:35:91:c4:5d:ec:8d:d6:3d:d9:72:ff:07:51:ea:d7:e4:a5:15'. There is a 'Create key pair' button in the top right of the table area.

Click on “create security group”

The screenshot shows the 'Create security group' page. The title is 'Create security group' with an 'Info' link. Below it is a note: 'A security group acts as a virtual firewall for your instance to control inbound and outbound traffic. To create a new security group, complete the fields below.' The 'Basic details' section contains fields for 'Security group name' (set to 'newsig'), 'Description' (set to 'my sgroup for my instances'), and 'VPC' (set to 'vpc-0fc65fb524e2580d1'). The 'Inbound rules' section has a note: 'This security group has no inbound rules.' At the bottom, there are 'Feedback' and 'English (US)' buttons, and a footer with copyright information and links to 'Privacy Policy' and 'Terms of Use'.

Add rules for ‘Inbound rules’

The screenshot shows the 'Create Security Group' page in the AWS Management Console. The 'Inbound rules' section is visible, displaying a list of predefined rules:

Type	Protocol	Port range	Source	Description
All TCP	TCP	0 - 65535	Anywhere	0.0.0.0/0
Custom ICMP - IPv4	All	All	Anywhere	0.0.0.0/0
SSH	TCP	22	Anywhere	0.0.0.0/0
HTTP	TCP	80	Anywhere	0.0.0.0/0
SMTP	TCP	25	Anywhere	0.0.0.0/0
HTTPS	TCP	443	Anywhere	0.0.0.0/0

An 'Add rule' button is located at the bottom left of the Inbound rules section.

Here I have specified 'anywhere' for source which should be changed to 'my ip', 'security group' or other specifications.

Add rules for outbound

The screenshot shows the 'Create Security Group' page in the AWS Management Console. The 'Outbound rules' section is visible, displaying a list of predefined rules:

Type	Protocol	Port range	Destination	Description
All TCP	TCP	0 - 65535	Anywhere	0.0.0.0/0
All ICMP - IPv4	ICMP	All	Anywhere	0.0.0.0/0
SSH	TCP	22	Anywhere	0.0.0.0/0
HTTP	TCP	80	Anywhere	0.0.0.0/0
HTTPS	TCP	443	Anywhere	0.0.0.0/0
SMTP	TCP	25	Anywhere	0.0.0.0/0

An 'Add rule' button is located at the bottom left of the Outbound rules section. A 'Create security group' button is at the bottom right.

Now let's create instances

The screenshot shows the AWS Management Console with the EC2 Management section selected. A modal window is open for creating a new security group. The modal contains fields for 'VPC ID' (set to vpc-0fc65fb524e2580d1) and 'In - optional' (with a link to 'Edit inbound rules').

Click on Launch Instance

The screenshot shows the AWS Management Console with the EC2 Dashboard selected. The 'Launch instance' button is highlighted in red, indicating it is the target for the click action.

Step 1: Choose an Amazon Machine Image (AMI)

Ubuntu Server 16.04 LTS (HVM, SSD Volume Type) - ami-085925c297f89fce1 (64-bit x86) / ami-05d7ab19b28efaf213 (64-bit Arm)

Free tier eligible

Root device type: hvm Virtualization type: hvm ENA Enabled: Yes

Are you launching a database instance? Try Amazon RDS.

Amazon Relational Database Service (RDS) makes it easy to set up, operate, and scale your database on AWS by automating time-consuming database management tasks. With RDS, you can easily deploy Amazon Aurora, MariaDB, MySQL, Oracle, PostgreSQL, and SQL Server databases on AWS. Aurora is a MySQL- and PostgreSQL-compatible, enterprise-class database at 1/10th the cost of commercial databases. Learn more about RDS

Select

Ubuntu Server 16.04 LTS (HVM, EBS General Purpose (SSD) Volume Type. Support available from Canonical (<http://www.ubuntu.com/cloud/services>))

Free tier eligible

Root device type: hvm Virtualization type: hvm ENA Enabled: Yes

Microsoft Windows Server 2019 Base - ami-0c278895328cdffdd

Free tier eligible

Root device type: hvm Virtualization type: hvm ENA Enabled: Yes

Deep Learning AMI (Ubuntu 18.04) Version 28.0 - ami-0c5642bef8d041bb

MXNet-1.6.0, Tensorflow-2.1.0 & 1.15.2, PyTorch-1.4.0, Keras-2.2, & other frameworks, configured with Neuron, NVIDIA CUDA, cuDNN, NCCL, Intel MKL-DNN, Docker & NVIDIA-Docker. For fully managed experience, check: <https://aws.amazon.com/sagemaker>

Root device type: hba Virtualization type: hvm ENA Enabled: Yes

Deep Learning AMI (Ubuntu 16.04) Version 28.0 - ami-05f3033402be9c7cd

Select

Select

Select

Select

Select

Feedback English (US) © 2008 - 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved. Privacy Policy Terms of Use

Choose Ubuntu Server 16.04 (AMI)

This is eligible for free tier

Step 2: Choose an Instance Type

Amazon EC2 provides a wide selection of instance types optimized to fit different use cases. Instances are virtual servers that can run applications. They have varying combinations of CPU, memory, storage, and networking capacity, and give you the flexibility to choose the appropriate mix of resources for your applications. Learn more about instance types and how they can meet your computing needs.

Currently selected: t2.micro (Variable ECUs, 1 vCPUs, 2.5 GHz, Intel Xeon Family, 1 GiB memory, EBS only)									
	Family	Type	vCPUs	Memory (GiB)	Instance Storage (GB)	EBS Optimized Available	Network Performance	IPv6 Support	
<input checked="" type="checkbox"/>	General purpose	t2.nano	1	0.5	EBS only	-	Low to Moderate	Yes	
<input checked="" type="checkbox"/>	General purpose	t2.micro	1	1	EBS only	-	Low to Moderate	Yes	Free tier eligible
<input type="checkbox"/>	General purpose	t2.small	1	2	EBS only	-	Low to Moderate	Yes	
<input type="checkbox"/>	General purpose	t2.medium	2	4	EBS only	-	Low to Moderate	Yes	
<input type="checkbox"/>	General purpose	t2.large	2	8	EBS only	-	Low to Moderate	Yes	
<input type="checkbox"/>	General purpose	t2.xlarge	4	16	EBS only	-	Moderate	Yes	
<input type="checkbox"/>	General purpose	t2.2xlarge	8	32	EBS only	-	Moderate	Yes	
<input type="checkbox"/>	General purpose	t3a.nano	2	0.5	EBS only	Yes	Up to 5 Gigabit	Yes	
<input type="checkbox"/>	General purpose	t3a.micro	2	1	EBS only	Yes	Up to 5 Gigabit	Yes	
<input type="checkbox"/>	General purpose	t3a.small	2	2	EBS only	Yes	Up to 5 Gigabit	Yes	

Cancel Previous Review and Launch Next: Configure Instance Details

Feedback English (US) © 2008 - 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved. Privacy Policy Terms of Use

However we need powerful machines to setup a Cloudera Hadoop Cluster

Thus, let's choose t2.large

Step 2: Choose an Instance Type

Amazon EC2 provides a wide selection of instance types optimized to fit different use cases. Instances are virtual servers that can run applications. They have varying combinations of CPU, memory, storage, and networking capacity, and give you the flexibility to choose the appropriate mix of resources for your applications. [Learn more](#) about instance types and how they can meet your computing needs.

Family	Type	vCPUs	Memory (GiB)	Instance Storage (GB)	EBS Optimized Available	Network Performance	IPv6 Support
General purpose	t2.nano	1	0.5	EBS only	-	Low to Moderate	Yes
General purpose	t2.micro	1	1	EBS only	-	Low to Moderate	Yes
General purpose	t2.small	1	2	EBS only	-	Low to Moderate	Yes
General purpose	t2.medium	2	4	EBS only	-	Low to Moderate	Yes
General purpose	t2.large	2	8	EBS only	-	Low to Moderate	Yes
General purpose	t2.xlarge	4	16	EBS only	-	Moderate	Yes
General purpose	t2.2xlarge	8	32	EBS only	-	Moderate	Yes
General purpose	t3a.nano	2	0.5	EBS only	Yes	Up to 5 Gigabit	Yes
General purpose	t3a.micro	2	1	EBS only	Yes	Up to 5 Gigabit	Yes
General purpose	t3a.small	2	2	EBS only	Yes	Up to 5 Gigabit	Yes

Review and Launch

Click on “configure instance details”

Provide details as shown

Step 3: Configure Instance Details

Number of instances: 3

Purchasing option: Request Spot instances

Network: vpc-0f65fb524e2580d1 (default) | Create new VPC

Subnet: No preference (default subnet in any Availability Zone) | Create new subnet

Auto-assign Public IP: Enable

Placement group: Add instance to placement group

Capacity Reservation: Open | Create new Capacity Reservation

IAM role: None | Create new IAM role

Shutdown behavior: Stop

Stop - Hibernate behavior: Enable hibernation as an additional stop behavior

Enable termination protection: Protect against accidental termination

Monitoring: Enable CloudWatch detailed monitoring
Additional charges apply

Tenancy: Shared - Run a shared hardware instance

Review and Launch

Click on Add storage

Configure as shown

Step 4: Add Storage

Your instance will be launched with the following storage device settings. You can attach additional EBS volumes and instance store volumes to your instance, or edit the settings of the root volume. You can also attach additional EBS volumes after launching an instance, but not instance store volumes. [Learn more about storage options in Amazon EC2.](#)

Volume Type	Device	Snapshot	Size (GiB)	Volume Type	IOPS	Throughput (MB/s)	Delete on Termination	Encryption
Root	/dev/sda1	snap-05945217d0601a348	8	General Purpose SSD (gp2)	N/A	100 / 3000	<input checked="" type="checkbox"/>	Not Encrypted

Add New Volume

Free tier eligible customers can get up to 30 GB of EBS General Purpose (SSD) or Magnetic storage. [Learn more about free usage tier eligibility and usage restrictions.](#)

Cancel Previous Review and Launch Next: Add Tags

Note ** We can use additional Volumes by clicking on “Add volume” and choose EBS, but as of now that is not required.

Click on “Add tags” [Nothing required here as of now]

Click on “Configure Security Group”

Step 6: Configure Security Group

A security group is a set of firewall rules that control the traffic for your instance. On this page, you can add rules to allow specific traffic to reach your instance. For example, if you want to set up a web server and allow internet traffic to reach your instance, add rules that allow unrestricted access to the HTTP and HTTPS ports. You can create a new security group or select from an existing one below. [Learn more about Amazon EC2 security groups.](#)

Assign a security group:

- Create a new security group
- Select an existing security group

Security Group ID	Name	Description	Actions
sg-0f5bedbe5c2bcbbbea	default	default VPC security group	Copy to new
sg-03cad200d510b25b	newsg	my sgroup for my instances	Copy to new

Select a security group above to view its inbound rules.

Cancel Previous Review and Launch

Feedback English (US) © 2008 - 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved. Privacy Policy Terms of Use

Step 6: Configure Security Group

A security group is a set of firewall rules that control the traffic for your instance. On this page, you can add rules to allow specific traffic to reach your instance. For example, if you want to set up a web server and allow Internet traffic to reach your instance, add rules that allow unrestricted access to the HTTP and HTTPS ports. You can create a new security group or select from an existing one below. [Learn more](#) about Amazon EC2 security groups.

Assign a security group: Create a new security group
 Select an existing security group

Security Group ID	Name	Description	Actions
sg_03dcad200d510b25b	newsg	my sgroup for my instances	Copy to new Copy to new

Inbound rules for sg_03dcad200d510b25b (Selected security groups: sg_03dcad200d510b25b)

Type	Protocol	Port Range	Source	Description
HTTP	TCP	80	0.0.0.0/0	
HTTP	TCP	80	0.0.0.0/0	
All TCP	TCP	0 - 65535	0.0.0.0/0	
All TCP	TCP	0 - 65535	0.0.0.0/0	
SSH	TCP	22	0.0.0.0/0	

[Cancel](#) [Previous](#) [Review and Launch](#)

After selecting your already created security group,

Click on “review and launch”

Step 7: Review Instance Launch

Please review your instance launch details. You can go back to edit changes for each section. Click **Launch** to assign a key pair to your instance and complete the launch process.

⚠ Your instance configuration is not eligible for the free usage tier
 To launch an instance that's eligible for the free usage tier, check your AMI selection, instance type, configuration options, or storage devices. Learn more about [free usage tier](#) eligibility and usage restrictions.
[Don't show me this again](#)

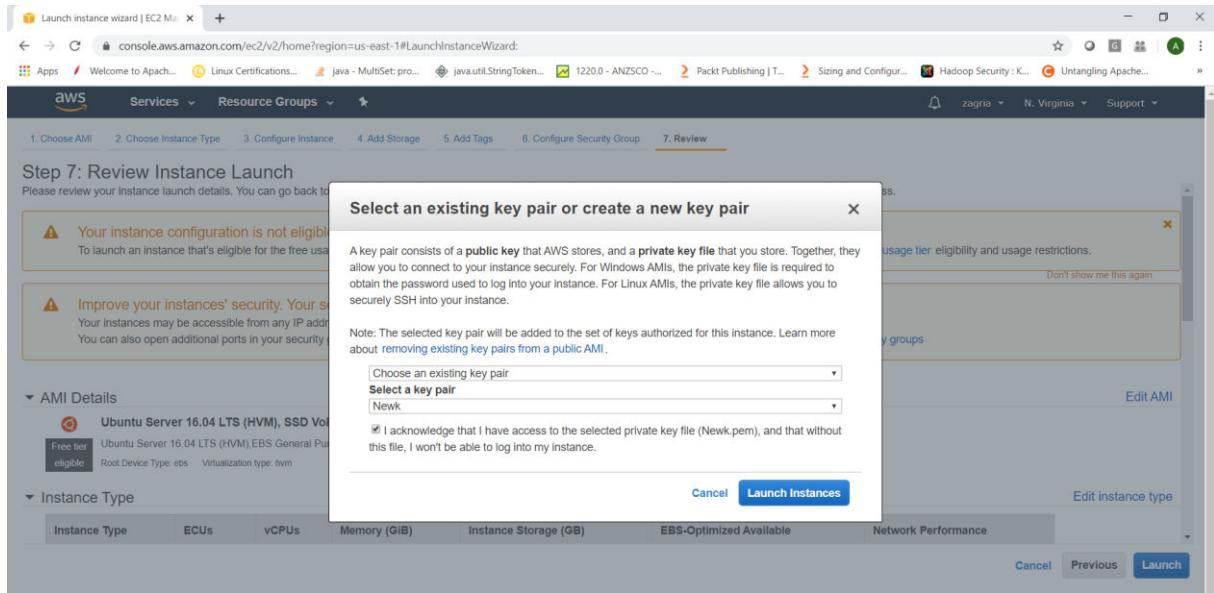
⚠ Improve your instances' security. Your security group, newsg, is open to the world.
 Your instances may be accessible from any IP address. We recommend that you update your security group rules to allow access from known IP addresses only.
 You can also open additional ports in your security group to facilitate access to the application or service you're running, e.g., HTTP (80) for web servers. [Edit security groups](#)

AMI Details
 Ubuntu Server 16.04 LTS (HVM), SSD Volume Type - ami-039a49e70ea773ff
 Free tier eligible
 Root Device Type: ebs Virtualization type: hvm [Edit AMI](#)

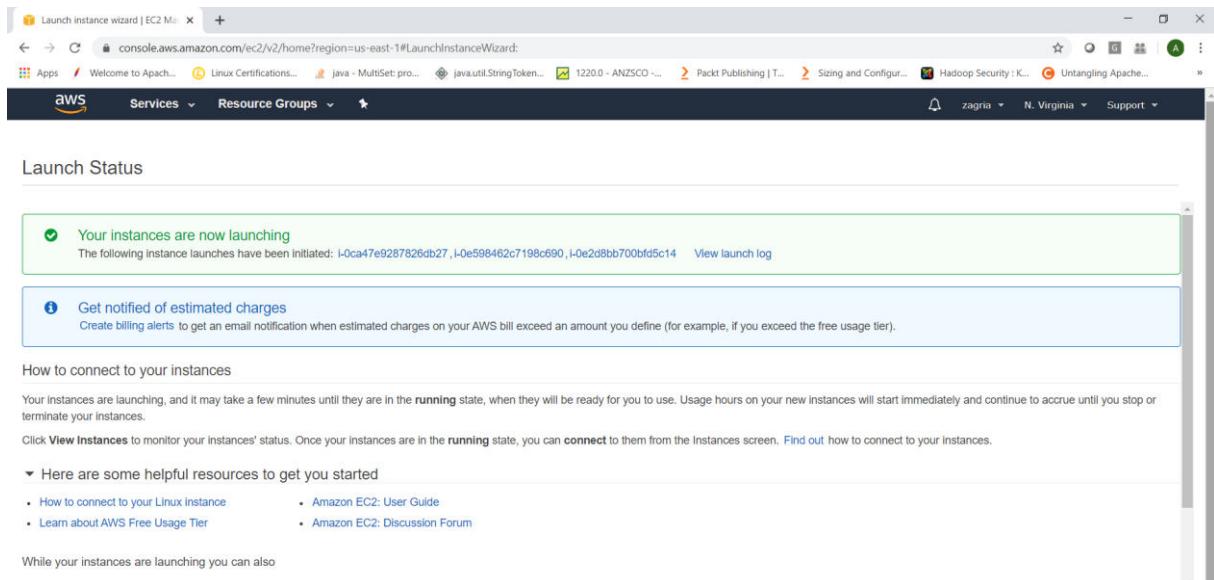
Instance Type
 Edit instance type
 Instance Type ECUs vCPUs Memory (GiB) Instance Storage (GB) EBS-Optimized Available Network Performance
[Cancel](#) [Previous](#) [Launch](#)

Click on “Launch”

Select your key pair



& “Launch Instances”



Click on any instance id,

The screenshot shows the AWS EC2 Management Console. In the left sidebar, under 'INSTANCES', 'Instances' is selected. In the main content area, a search bar at the top contains the text 'search : i-0e2d8bb700bfd5c14'. Below it is a table with columns: Name, Instance ID, Instance Type, Availability Zone, Instance State, Status Checks, Alarm Status, Public DNS (IPv4), IPv4 Public IP, and IPv6 Public IP. One row is highlighted, corresponding to the search term. A detailed view of this instance is shown in a modal window below.

Name	Instance ID	Instance Type	Availability Zone	Instance State	Status Checks	Alarm Status	Public DNS (IPv4)	IPv4 Public IP	IPv6 Public IP
	i-0e2d8bb700bfd5c14	t2.large	us-east-1b	running	Initializing	Loading...	ec2-3-83-50-248.compute-1.amazonaws.com	3.83.50.248	-

Instance: i-0e2d8bb700bfd5c14 Public DNS: ec2-3-83-50-248.compute-1.amazonaws.com

Description **Status Checks** **Monitoring** **Tags**

Instance ID	i-0e2d8bb700bfd5c14	Public DNS (IPv4)	ec2-3-83-50-248.compute-1.amazonaws.com
Instance state	running	IPv4 Public IP	3.83.50.248
Instance type	t2.large	IPv6 IPs	-
Finding	Opt-in to AWS Compute Optimizer for recommendations. Learn more	Elastic IPs	
Private DNS	ip-172-31-90-165.ec2.internal	Availability zone	us-east-1b
Private IPs	172.31.90.165	Security groups	newsg, view inbound rules , view outbound rules
Secondary private IPs		Scheduled events	No scheduled events
VPC ID	vpc-0fc65fb524e2580d1	AMI ID	Loading ami-039a49e70ea773ffc...
Subnet ID	subnet-00-20860a7-r149r	Platform details	loading

& in search ,remove the instance ID from search

That should show you all your instances

The screenshot shows the AWS EC2 Management Console with a search bar containing 'Filter by tags and attributes or search by keyword'. Below it is a table with the same columns as the previous screenshot. Three instances are listed, each with a different Public DNS and IPv4 address. A detailed view of the first instance is shown in a modal window below.

Name	Instance ID	Instance Type	Availability Zone	Instance State	Status Checks	Alarm Status	Public DNS (IPv4)	IPv4 Public IP	IPv6 Public IP
	i-0ca47e9287826db27	t2.large	us-east-1b	running	Initializing	None	ec2-3-84-220-90.compute-1.amazonaws.com	3.84.220.90	-
	i-0e2d8bb700bfd5c14	t2.large	us-east-1b	running	Initializing	None	ec2-3-83-50-248.compute-1.amazonaws.com	3.83.50.248	-
	i-0e598462c7198c690	t2.large	us-east-1b	running	Initializing	None	ec2-3-84-101-212.compute-1.amazonaws.com	3.84.101.212	-

Instance: i-0e2d8bb700bfd5c14 Public DNS: ec2-3-83-50-248.compute-1.amazonaws.com

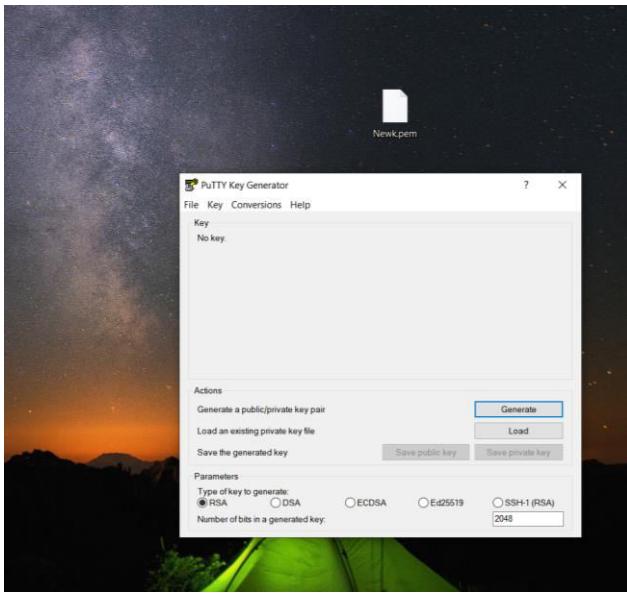
Description **Status Checks** **Monitoring** **Tags**

Instance ID	i-0e2d8bb700bfd5c14	Public DNS (IPv4)	ec2-3-83-50-248.compute-1.amazonaws.com
Instance state	running	IPv4 Public IP	3.83.50.248
Instance type	t2.large	IPv6 IPs	-
Finding	Opt-in to AWS Compute Optimizer for recommendations. Learn more	Elastic IPs	
Private DNS	ip-172-31-90-165.ec2.internal	Availability zone	us-east-1b
Private IPs	172.31.90.165	Security groups	newsg, view inbound rules , view outbound rules
Secondary private IPs		Scheduled events	No scheduled events
VPC ID	vpc-0fc65fb524e2580d1	AMI ID	ubuntu/images/hvm-ssd/ubuntu-xenial-16.04-amd64-server-20200407 (ami-039a49e70ea773ffc)

If All steps completed so far, your instances should be ready to be used.

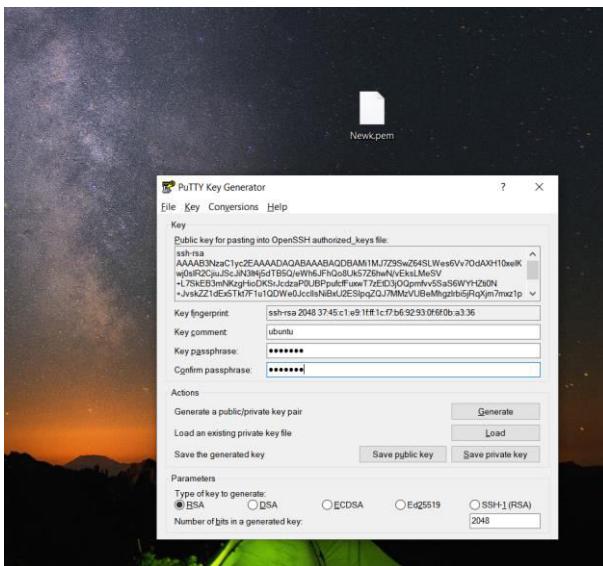
To connect these instances using external ssh client such as “putty”, we need to convert “*.pem” to “*.ppk” file for which we can use “puttygen.exe”

Launch puttygen

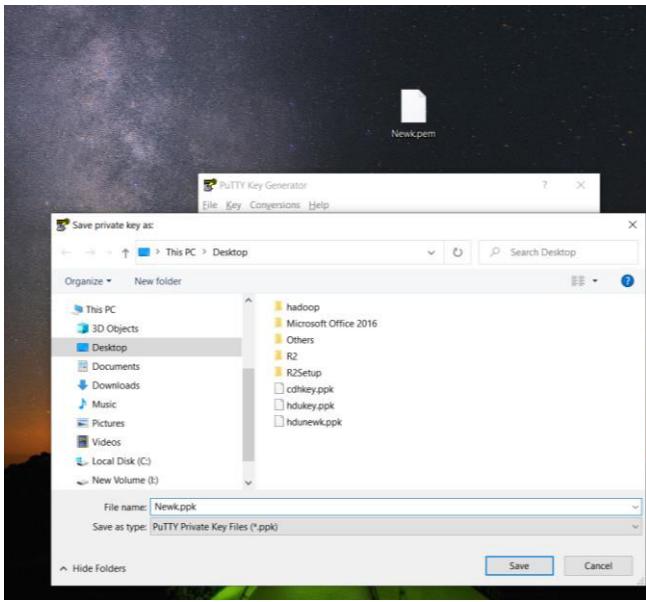


Click on “Load” and add your “*.pem” file.

Now give credentials



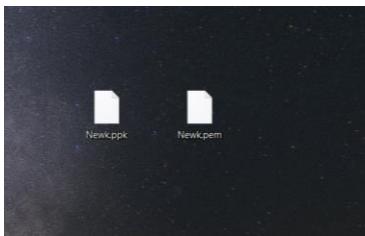
Click on “save private key” which should allow you to save your key as “*.ppk” file.



Now you have “*.pem” file to do ssh to your ec2 instances (say from one instance to another or when using some external software to ssh into your instances)

&

You have “*.ppk” file to ssh into your instances using external ssh client such as putty.



Now setup putty to connect to these instances.

Rename your instances to say “c1,c2,c3” for easier access.

Screenshot of the AWS EC2 Management Console showing two separate instances (c1 and c2) running in us-east-1b. The details for instance c1 are shown in the modal.

Name	Instance ID	Instance Type	Availability Zone	Instance State	Status Checks	Alarm Status	Public DNS (IPv4)	IPv4 Public IP	IPv6
c1	i-0ca47e9287826db27	t2 large	us-east-1b	running	2/2 checks ...	None	ec2-3-84-220-90.compute-1.amazonaws.com	3.84.220.90	-
2/255	i-0e598462c7198c690	t2 large	us-east-1b	running	2/2 checks ...	None	ec2-3-83-50-248.compute-1.amazonaws.com	3.83.50.248	-

Instance: i-0ca47e9287826db27 Public DNS: ec2-3-84-220-90.compute-1.amazonaws.com

Description

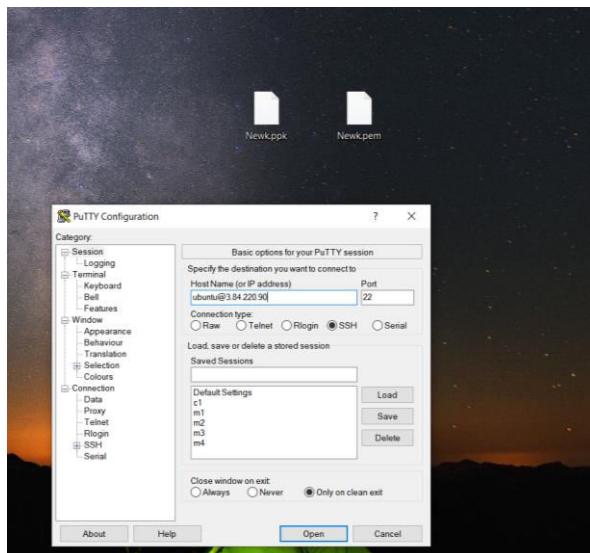
Instance ID	i-0ca47e9287826db27	Public DNS (IPv4)	ec2-3-84-220-90.compute-1.amazonaws.com
Instance state	running	IPv4 Public IP	3.84.220.90
Instance type	t2 large	IPv6 IPs	-
Finding	Opt-in to AWS Compute Optimizer for recommendations. Learn more	Elastic IPs	-
Private DNS	ip-172-31-95-168.ec2.internal	Availability zone	us-east-1b
Private IPs	172.31.95.168	Security groups	newsg, view inbound rules, view outbound rules
Secondary private IPs		Scheduled events	No scheduled events
VPC ID	vpc-0fc65fb524e2580d1	AMI ID	ubuntu/images/hvm-ssd/ubuntu-xenial-16.04-amd64-server.20200407 (ami-039a49e70ea773ffc)

Instance: i-0e598462c7198c690 (c3) Public DNS: ec2-3-84-101-212.compute-1.amazonaws.com

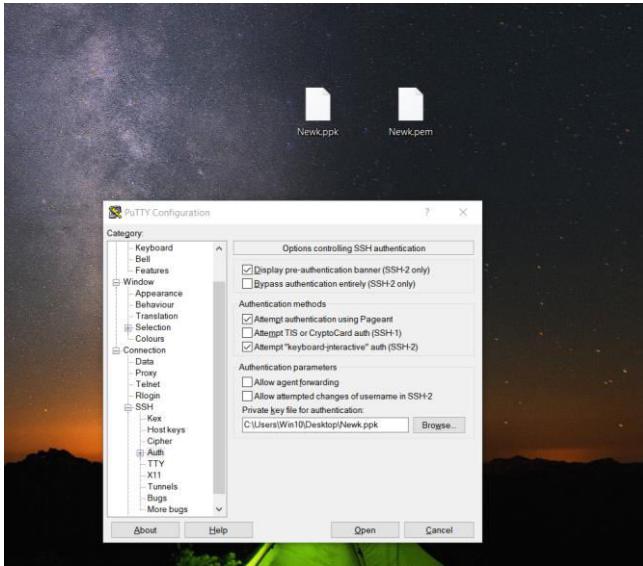
Description

Instance ID	i-0e598462c7198c690	Public DNS (IPv4)	ec2-3-84-101-212.compute-1.amazonaws.com
Instance state	running	IPv4 Public IP	3.84.101.212
Instance type	t2 large	IPv6 IPs	-
Finding	Opt-in to AWS Compute Optimizer for recommendations. Learn more	Elastic IPs	-
Private DNS	ip-172-31-82-22.ec2.internal	Availability zone	us-east-1b
Private IPs	172.31.82.22	Security groups	newsg, view inbound rules, view outbound rules
Secondary private IPs		Scheduled events	No scheduled events
VPC ID	vpc-0fc65fb524e2580d1	AMI ID	ubuntu/images/hvm-ssd/ubuntu-xenial-16.04-amd64-server.20200407 (ami-039a49e70ea773ffc)

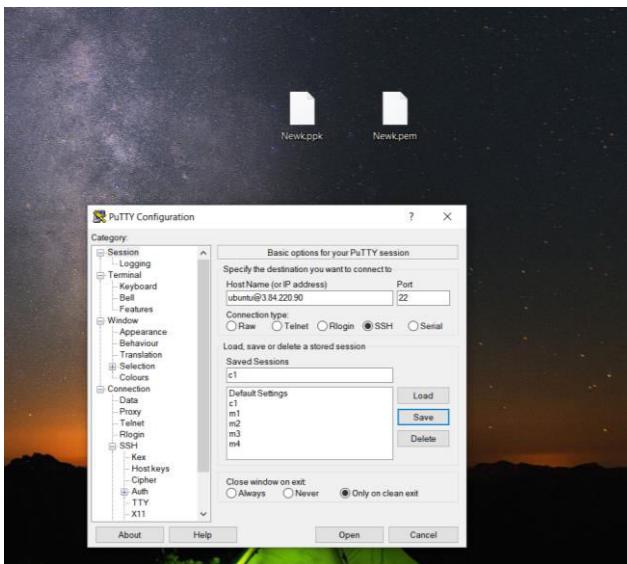
We need public IPs to connect from putty.



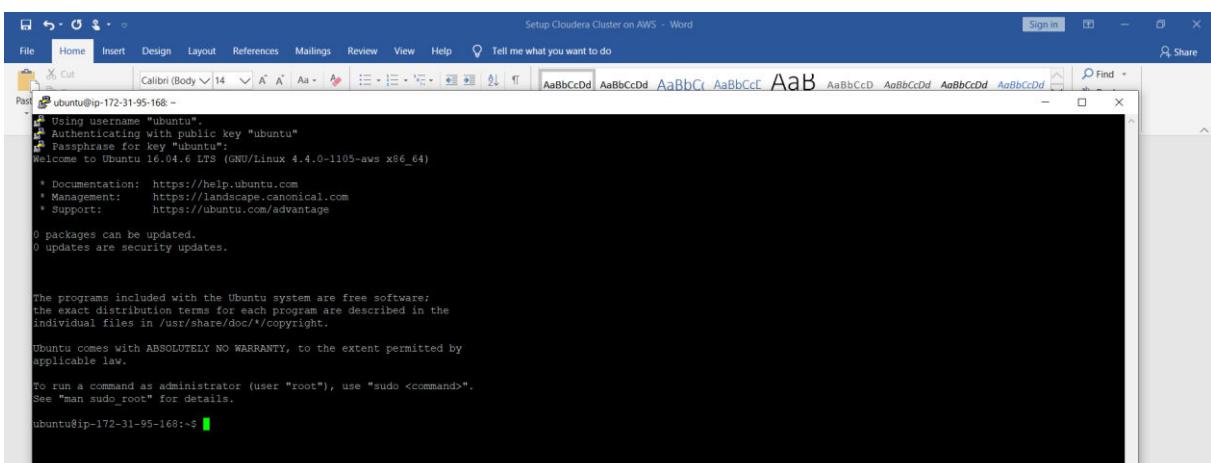
Click on ssh> auth and add your “*.ppk” file



Save the session as c1

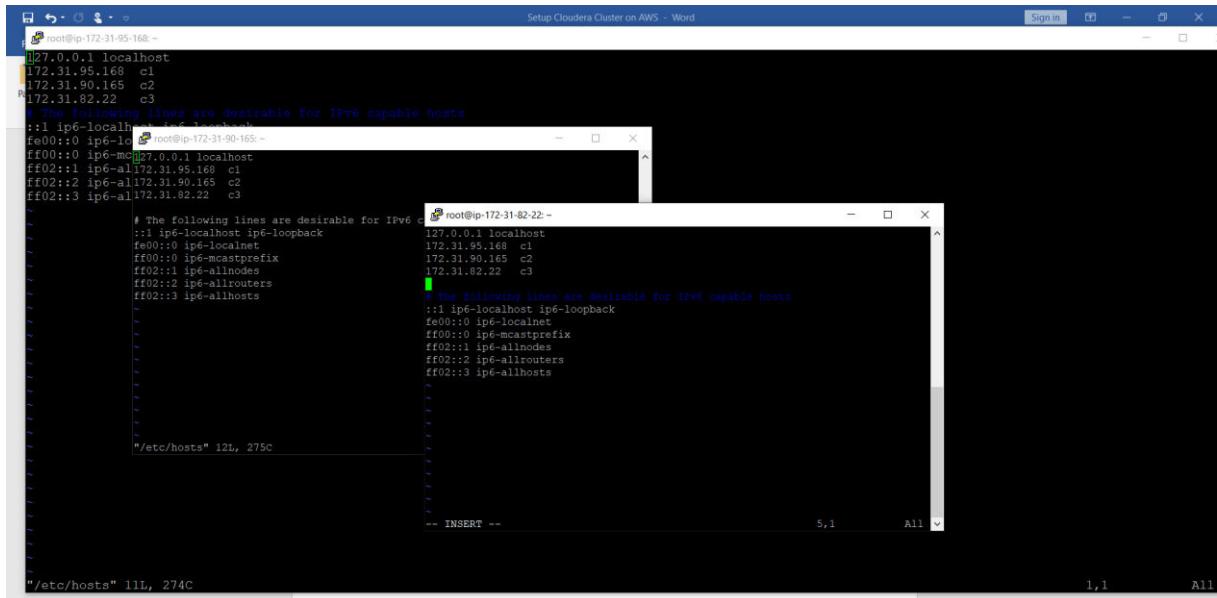


Click on open



Similarly you can add other machines.

Now update /etc/hosts in all instances



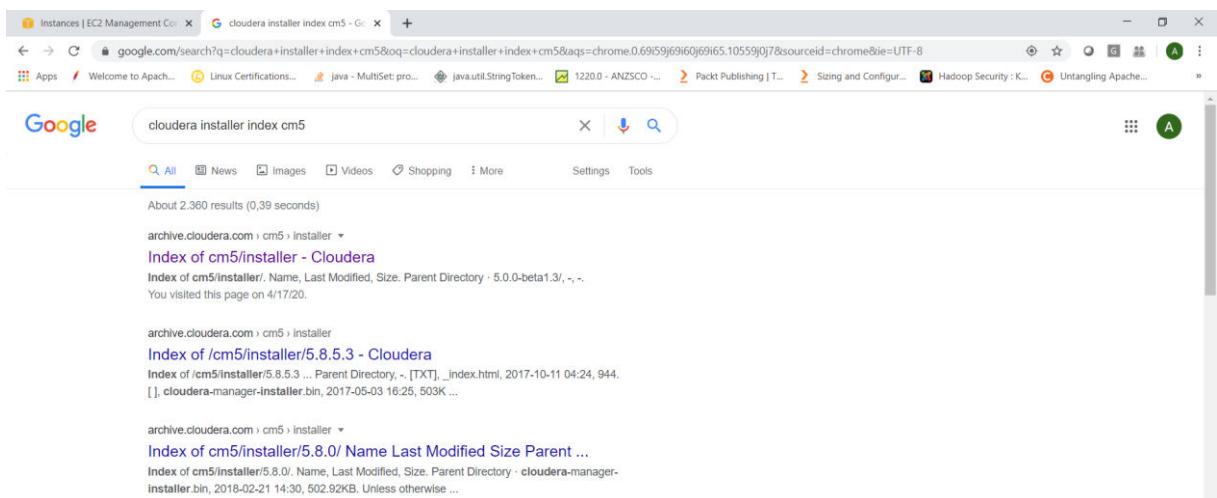
```
root@ip-172-31-95-168 ~
127.0.0.1 localhost
172.31.95.168 c1
172.31.90.165 c2
172.31.82.22 c3
# The following lines are desirable for IPv6 capable hosts
::1 ip6-localhost ip6-loopback
fe00::0 ip6-localnet
ff00::0 ip6-mcastprefix
ff02::1 ip6-allnodes
ff02::2 ip6-allrouters
ff02::3 ip6-allhosts

root@ip-172-31-82-22: ~
127.0.0.1 localhost
172.31.95.168 c1
172.31.90.165 c2
172.31.82.22 c3
# The following lines are desirable for IPv6 capable hosts
::1 ip6-localhost ip6-loopback
fe00::0 ip6-localnet
ff00::0 ip6-mcastprefix
ff02::1 ip6-allnodes
ff02::2 ip6-allrouters
ff02::3 ip6-allhosts

"/etc/hosts" 12L, 275C

-- INSERT --
5,1      All 1,1      All 1,1      All
```

Now we download the “cloudera installer bin file”



Google cloudera installer index cm5

About 2,360 results (0,39 seconds)

archive.cloudera.com › cm5 › installer

Index of cm5/installer - Cloudera

Index of cm5/installer/. Name, Last Modified, Size, Parent Directory · 5.0.0-beta1.3/, ~, ~.

You visited this page on 4/17/20.

archive.cloudera.com › cm5 › installer

Index of /cm5/installer/5.8.5.3 - Cloudera

Index of /cm5/installer/5.8.5.3 ... Parent Directory, ~, [TXT], _index.html, 2017-10-11 04:24, 944, [1], cloudera-manager-installer.bin, 2017-05-03 16:25, 509K ...

archive.cloudera.com › cm5 › installer

Index of cm5/installer/5.8.0/ Name Last Modified Size Parent ...

Index of cm5/installer/5.8.0/. Name, Last Modified, Size, Parent Directory · cloudera-manager-installer.bin, 2018-02-21 14:30, 502.92KB, Unless otherwise ...



Instances | EC2 Management Con... https://archive.cloudera.com/cm5/installer/5.16/ +

archive.cloudera.com/cm5/installer/5.16/

Index of cm5/installer/5.16/

Name	Last Modified	Size
Parent Directory		
cloudera-manager-installer.bin	2019-06-18 13:10	509.28KB

Unless otherwise specified herein, downloads of software from this site and its use are governed by the [Cloudera Standard License](#). By downloading or using this software from this site you agree to be bound by the [Cloudera Standard License](#). If you do not wish to be bound by these terms, then do not download or use the software from this site.

```

Instances | EC2 Management Con... https://archive.cloudera.com/cm5/installer/5.16/
Index of /cm5/installer/5.16/
root@ip-172-31-95-168:~# hostname
ip-172-31-95-168
root@ip-172-31-95-168:~# vi /etc/hosts
root@ip-172-31-95-168:~# wget https://archive.cloudera.com/cm5/installer/5.16/cloudera-manager-installer.bin
--2020-05-04 01:06:14-- https://archive.cloudera.com/cm5/installer/5.16/cloudera-manager-installer.bin
Resolving archive.cloudera.com (archive.cloudera.com)... 151.101.248.167
Connecting to archive.cloudera.com (archive.cloudera.com)|151.101.248.167|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 521501 (509K) [application/octet-stream]
License: If saving to: 'cloudera-manager-installer.bin'

cloudera-manager-installer.bi 100%[=====] 509.28K --.-KB/s in 0.01s
2020-05-04 01:06:14 (48.8 MB/s) - 'cloudera-manager-installer.bin' saved [521501/521501]
root@ip-172-31-95-168:~#

```

Change the permission of this installer to 777

```

File Home Insert Design Layout References Mailings Review View Help Tell me what you want to do
Setup Cloudera Cluster on AWS - Word
Sign in Share
AaBbCcDd Intense E... Find Replace Select Editing
Paste For ip-172-31-95-168 Clipboar...
root@ip-172-31-95-168:~# hostname
root@ip-172-31-95-168:~# vi /etc/hosts
root@ip-172-31-95-168:~# wget https://archive.cloudera.com/cm5/installer/5.16/cloudera-manager-installer.bin
--2020-05-04 01:06:14-- https://archive.cloudera.com/cm5/installer/5.16/cloudera-manager-installer.bin
Resolving archive.cloudera.com (archive.cloudera.com)... 151.101.248.167
Connecting to archive.cloudera.com (archive.cloudera.com)|151.101.248.167|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 521501 (509K) [application/octet-stream]
Saving to: 'cloudera-manager-installer.bin'

cloudera-manager-installer.bi 100%[=====] 509.28K --.-KB/s in 0.01s
2020-05-04 01:06:14 (48.8 MB/s) - 'cloudera-manager-installer.bin' saved [521501/521501]
root@ip-172-31-95-168:~# chmod -R 777 cloudera-manager-installer.bin
root@ip-172-31-95-168:~# ls
cloudera-manager-installer.bin snap
root@ip-172-31-95-168:~#

```

Run the installer which should start “CDH” installation.

```

File Home Insert Design Layout References Mailings Review View Help Tell me what you want to do
Setup Cloudera Cluster on AWS - Word
Sign in Share
AaBbCcDd Intense E... Find Replace Select Editing
Paste For ip-172-31-95-168 Clipboar...
root@ip-172-31-95-168:~# hostname
root@ip-172-31-95-168:~# vi /etc/hosts
root@ip-172-31-95-168:~# wget https://archive.cloudera.com/cm5/installer/5.16/cloudera-manager-installer.bin
--2020-05-04 01:06:14-- https://archive.cloudera.com/cm5/installer/5.16/cloudera-manager-installer.bin
Resolving archive.cloudera.com (archive.cloudera.com)... 151.101.248.167
Connecting to archive.cloudera.com (archive.cloudera.com)|151.101.248.167|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 521501 (509K) [application/octet-stream]
Saving to: 'cloudera-manager-installer.bin'

cloudera-manager-installer.bi 100%[=====] 509.28K --.-KB/s in 0.01s
2020-05-04 01:06:14 (48.8 MB/s) - 'cloudera-manager-installer.bin' saved [521501/521501]
root@ip-172-31-95-168:~# chmod -R 777 cloudera-manager-installer.bin
root@ip-172-31-95-168:~# ls
cloudera-manager-installer.bin snap
root@ip-172-31-95-168:~# ./cloudera-manager-installer.bin

```

```
root@ip-172-31-95-168: ~
Cloudera Manager 5
Cloudera Standard License
Version 2018-08-14

THE TERMS AND CONDITIONS OF THIS CLOUDERA STANDARD LICENSE (THE "AGREEMENT") APPLY TO YOUR USE OF OR ACCESS TO THE PRODUCTS (AS DEFINED BELOW) MADE AVAILABLE BY CLOUDERA, INC. ("CLOUDERA").

PLEASE READ THIS AGREEMENT CAREFULLY.

IF YOU ("YOU" OR "CUSTOMER") PLAN TO USE OR ACCESS ANY OF THE PRODUCTS ON BEHALF OF A COMPANY OR OTHER ENTITY, YOU REPRESENT THAT YOU ARE THE EMPLOYEE OR AGENT OF SUCH COMPANY OR OTHER ENTITY AND YOU HAVE THE AUTHORITY TO ACCEPT ALL OF THE TERMS AND CONDITIONS SET FORTH IN THIS AGREEMENT ON BEHALF OF SUCH COMPANY OR OTHER ENTITY.

BY USING OR ACCESSING ANY OF THE PRODUCTS, YOU ACKNOWLEDGE AND AGREE THAT:
(A) YOU HAVE READ ALL OF THE TERMS OF THIS AGREEMENT;
(B) YOU UNDERSTAND ALL OF THE TERMS OF THIS AGREEMENT;
(C) YOU AGREE TO BE LEGALLY BOUND BY ALL OF THE TERMS SET FORTH IN THIS AGREEMENT.

IF YOU DO NOT AGREE WITH ANY OF THE TERMS OF THIS AGREEMENT, YOU MAY NOT USE OR ACCESS ANY PORTION OF THE PRODUCTS.

THE "EFFECTIVE DATE" OF THIS AGREEMENT IS THE DATE YOU FIRST DOWNLOAD OR ACCESS ANY OF THE PRODUCTS.

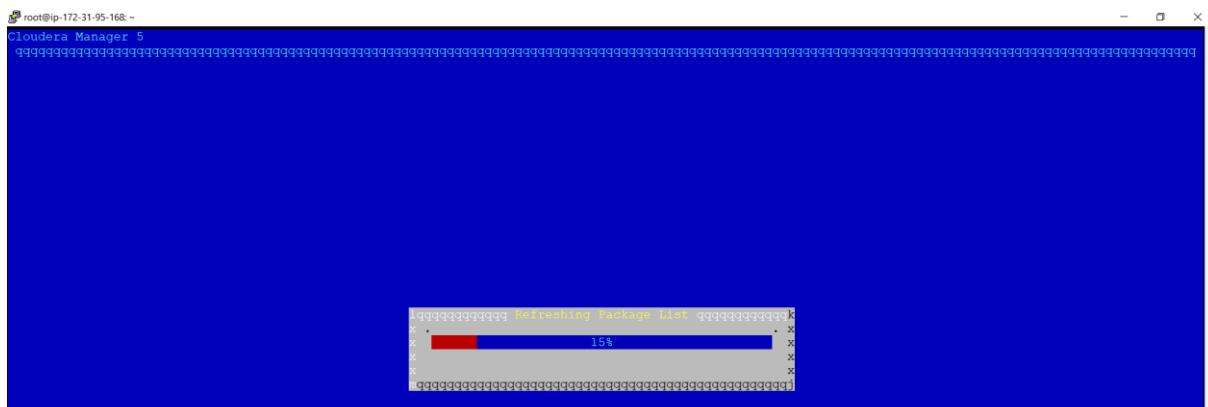
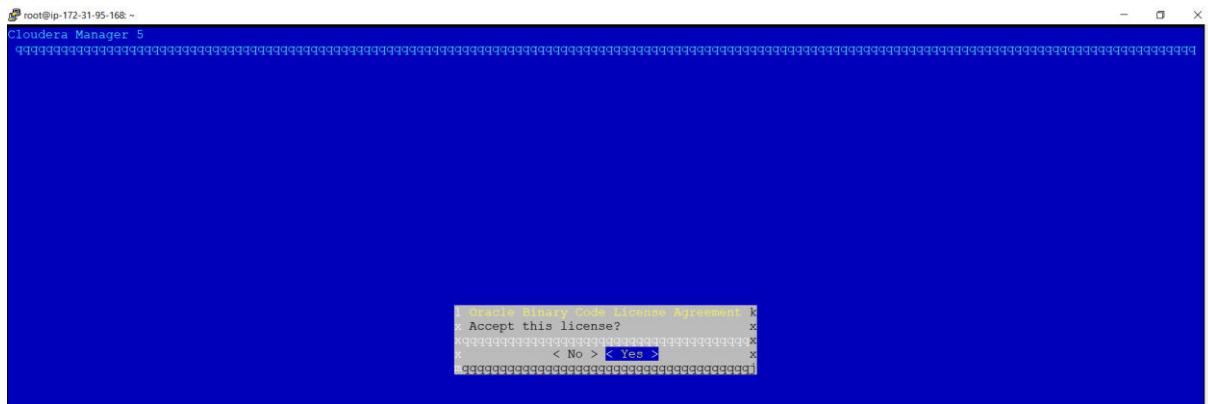
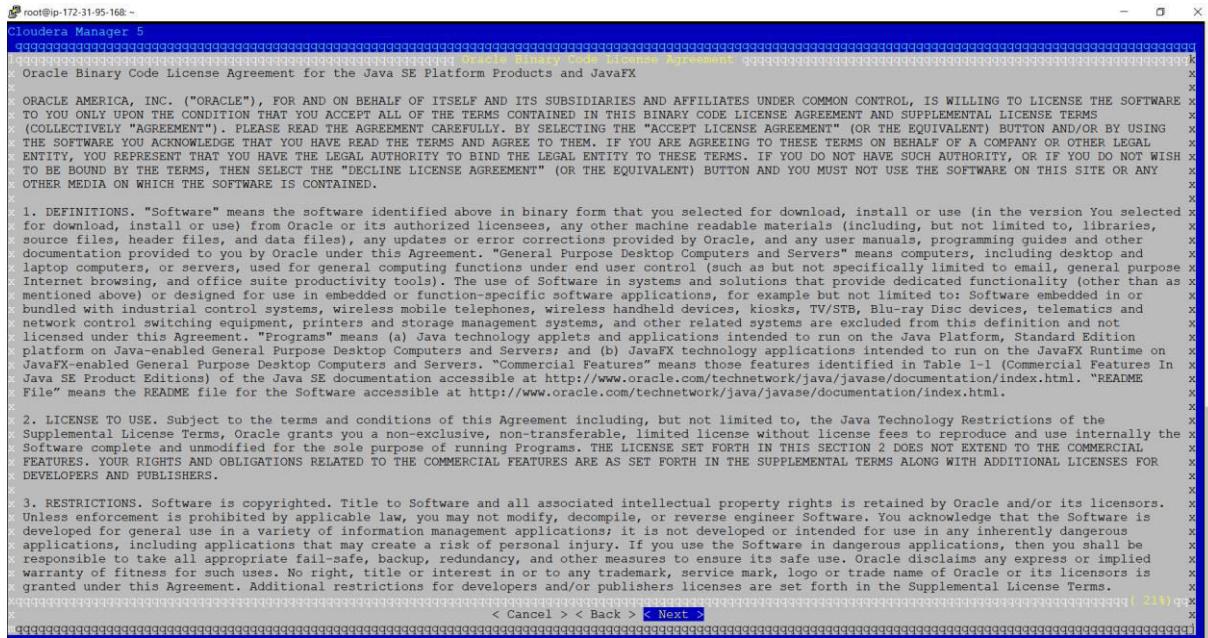
1. Product. For the purpose of this Agreement, "Product" shall mean any of Cloudera's offerings including but not limited to Cloudera proprietary software, any hosted or cloud-based service (a "Cloudera Online Service"), any trial software, and any software related to the foregoing.

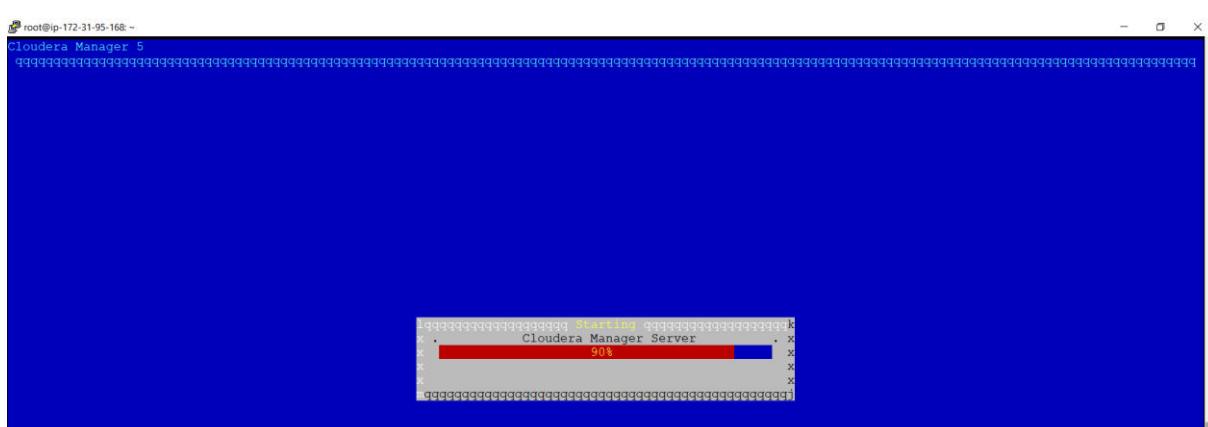
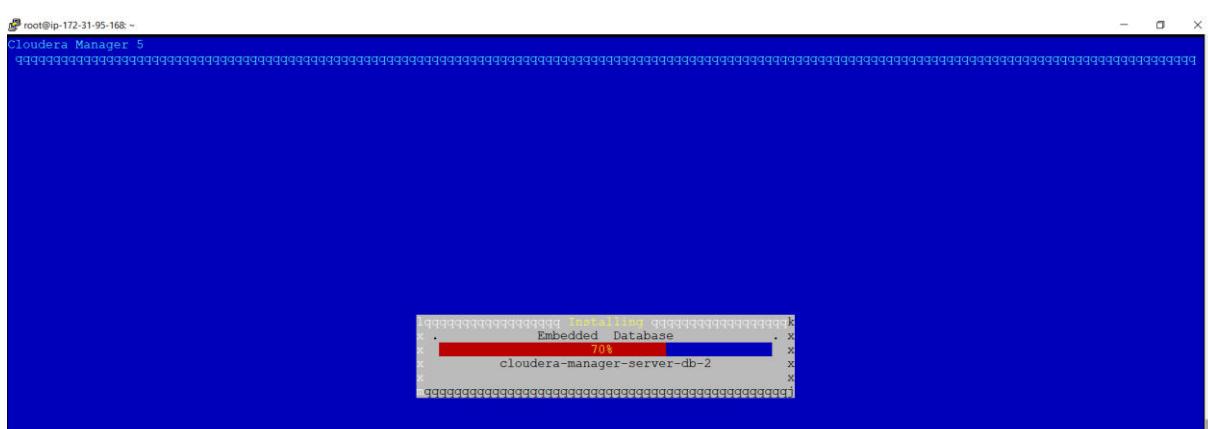
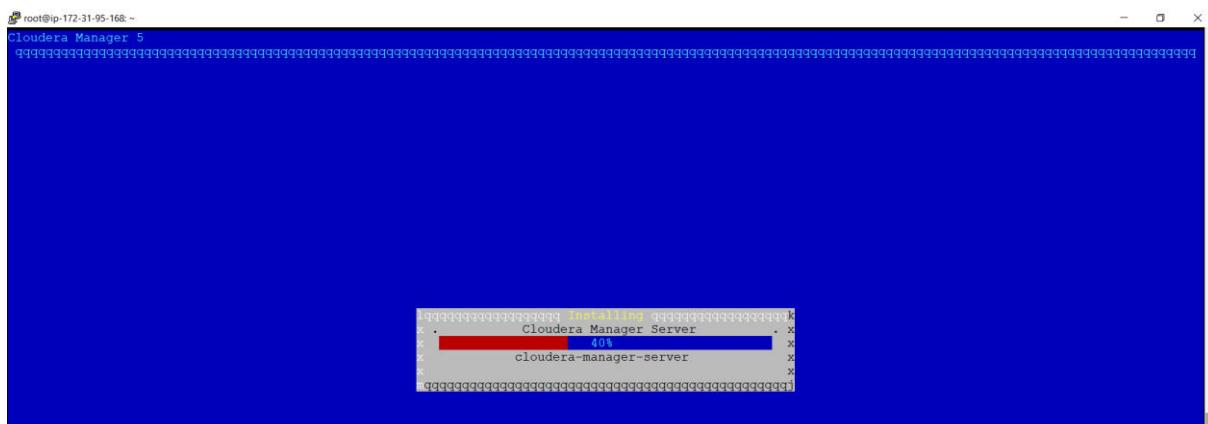
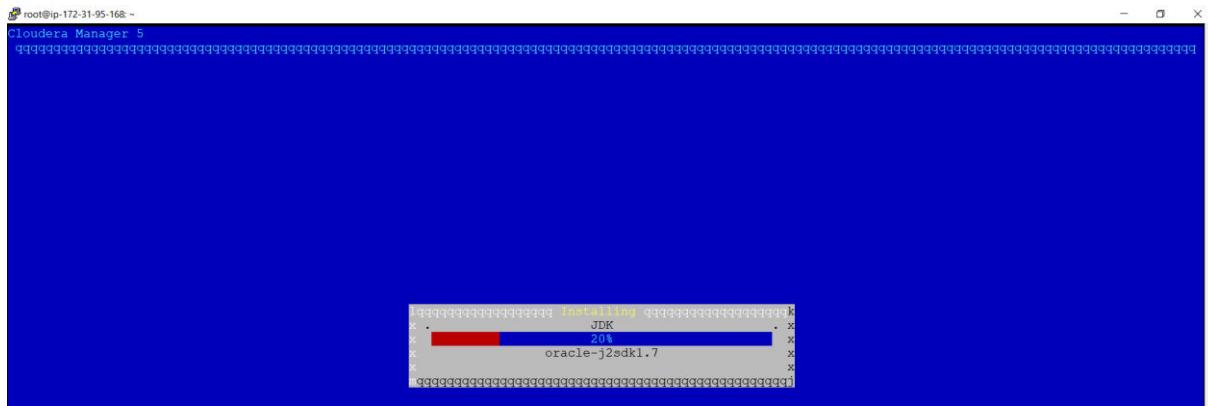
2. Entire Agreement. This Agreement includes any exhibits attached hereto and web links referenced herein or in any exhibit, and the terms set forth on the Cloudera web site at https://www.cloudera.com/documentation/other/licenses/topics/Third-Party-Licenses.html, all hereby incorporated by reference into this Agreement in their entirety as they appear on the Effective Date of this Agreement, and as may be updated by Cloudera in its sole discretion from time to time without amendment to this Agreement.

This Agreement is the entire agreement of the parties regarding the subject matter hereof, superseding all other agreements between the parties, whether oral or written, regarding the subject matter hereof.

3. License Delivery. Cloudera grants to Customer a nonexclusive, nontransferable, nonsublicensable, revocable and limited license to access and use the applicable Product(s) as defined above in Section 1 solely for Customer's internal purposes. The Product is delivered via electronic download (or, in the case of Cloudera Online Services, as a hosted, cloud-based service, accessible to Customer through a web browser), made available following Customer's acceptance of this Agreement. If Customer uses the Cloudera Online Services, Customer will adhere to and agrees to be bound by the Data Policy and the Privacy Policy, located at https://www.cloudera.com/legal/policies.html, as such policies may be updated by Cloudera from time to time (the "Data Policy" and the "Privacy Policy," respectively) without amendment to this Agreement.

< Cancel > < Back > < Next > ( 19 )
```





Waiting for few seconds or a minute before accessing the admin console.

A screenshot of a terminal window titled "root@ip-172-31-95-168: ~". The title bar also displays "Cloudera Manager 5". The main content of the terminal shows the following text:

```
[root@ip-172-31-95-168 ~]# ./cloudera-manager5.sh
[INFO] Cloudera Manager 5.12.0-1255
[INFO] Installation was successful.
[INFO] < OK >
```

The text is white on a black background, indicating a successful installation of Cloudera Manager 5.

At this time, check the status of “cloudera-scm-server”

```
root@ip-172-31-95-168:~# hostname
ip-172-31-95-168
root@ip-172-31-95-168:~# vi /etc/hosts
root@ip-172-31-95-168:~# vi /etc/hosts
root@ip-172-31-95-168:~# wget https://archive.cloudera.com/cm5/installer/5.16/cloudera-manager-installer.bin
2020-05-04 01:06:14-- https://archive.cloudera.com/cm5/installer/5.16/cloudera-manager-installer.bin
Resolving archive.cloudera.com (archive.cloudera.com)... 151.101.248.167
Connecting to archive.cloudera.com (archive.cloudera.com)|151.101.248.167|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 521501 (509K) [application/octet-stream]
Saving to: 'cloudera-manager-installer.bin'

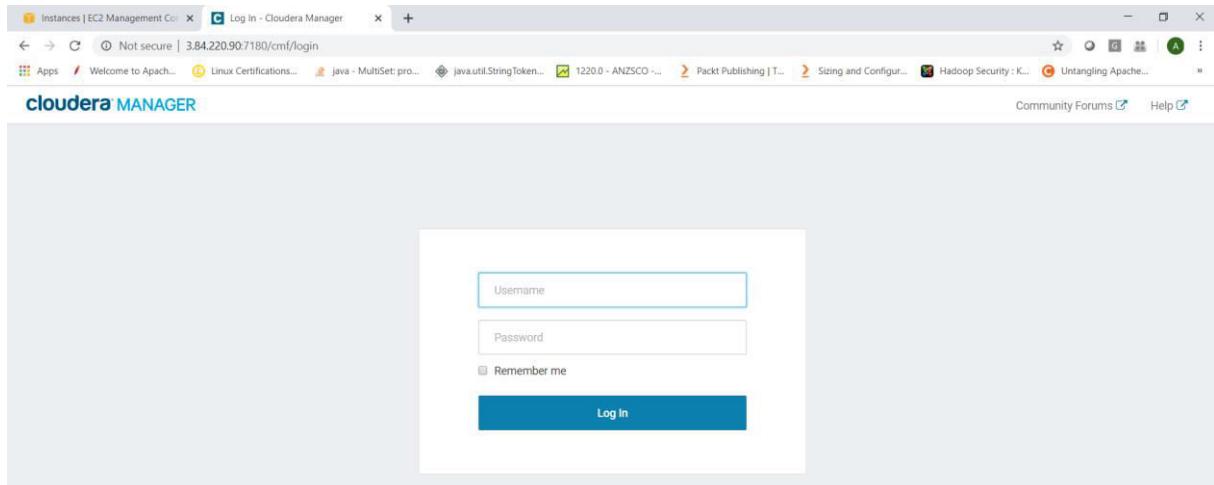
cloudera-manager-installer.bi 100%[=====] 509.28K --.-KB/s   in 0.01s

2020-05-04 01:06:14 (49.8 MB/s) - `cloudera-manager-installer.bin' saved [521501/521501]

root@ip-172-31-95-168:~# chmod -R 777 cloudera-manager-installer.bin
root@ip-172-31-95-168:~# ls
cloudera-manager-installer.bin  snap
root@ip-172-31-95-168:~# ./cloudera-manager-installer.bin
root@ip-172-31-95-168:~# service cloudera-scm-server status
● cloudera-scm-server.service - LSB: Cloudera SCM Server
   Loaded: loaded (/etc/init.d/cloudera-scm-server; bad; vendor preset: enabled)
   Active: active (exited) since Mon 2020-05-04 01:09:48 UTC; 1min 14s ago
     Docs: man:systemd-sysv-generator(8)
  Process: 14751 ExecStart=/etc/init.d/cloudera-scm-server start (code=exited, status=0/SUCCESS)

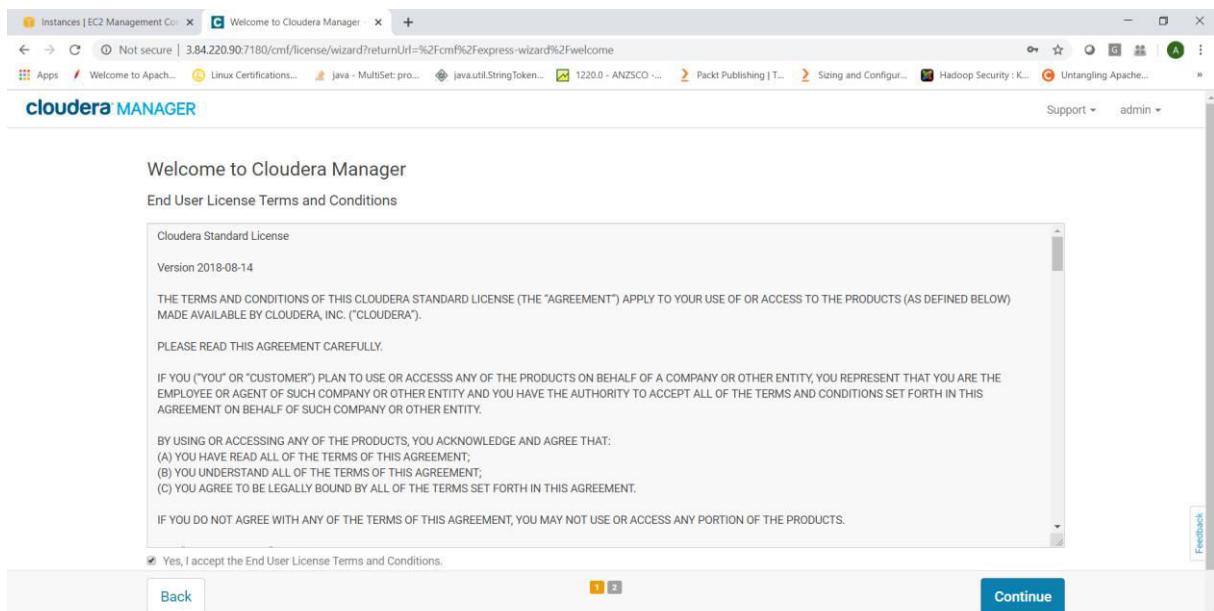
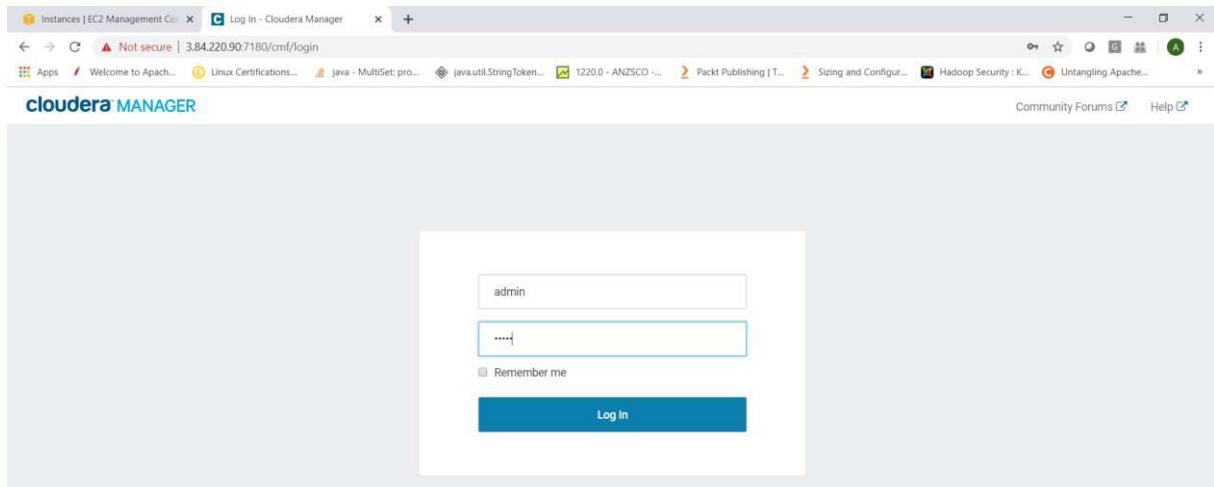
May 04 01:09:43 ip-172-31-95-168 systemd[1]: Starting LSB: Cloudera SCM Server...
May 04 01:09:43 ip-172-31-95-168 su[14773]: Successful su for cloudera-scm by root
May 04 01:09:43 ip-172-31-95-168 su[14773]: + ??? root:cloudera-scm
May 04 01:09:43 ip-172-31-95-168 su[14773]: pam_unix(su:session): session opened for user cloudera-scm by (uid=0)
May 04 01:09:48 ip-172-31-95-168 cloudera-scm-server[14751]: Starting cloudera-scm-server: * cloudera-scm-server started
May 04 01:09:48 ip-172-31-95-168 systemd[1]: Started LSB: Cloudera SCM Server.
root@ip-172-31-95-168:~#
```

Access your admin console by using public ip of the host where cloudera installer was run



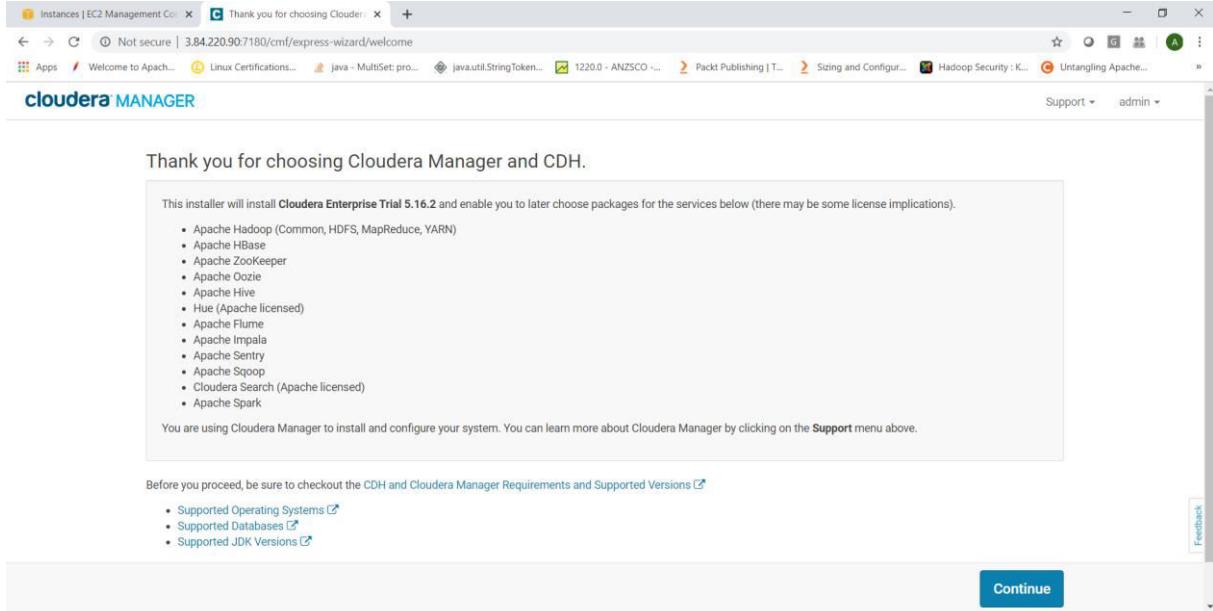
<http://publicipaddressofyourhost:7180>

Login with default id and pswd as 'admin'



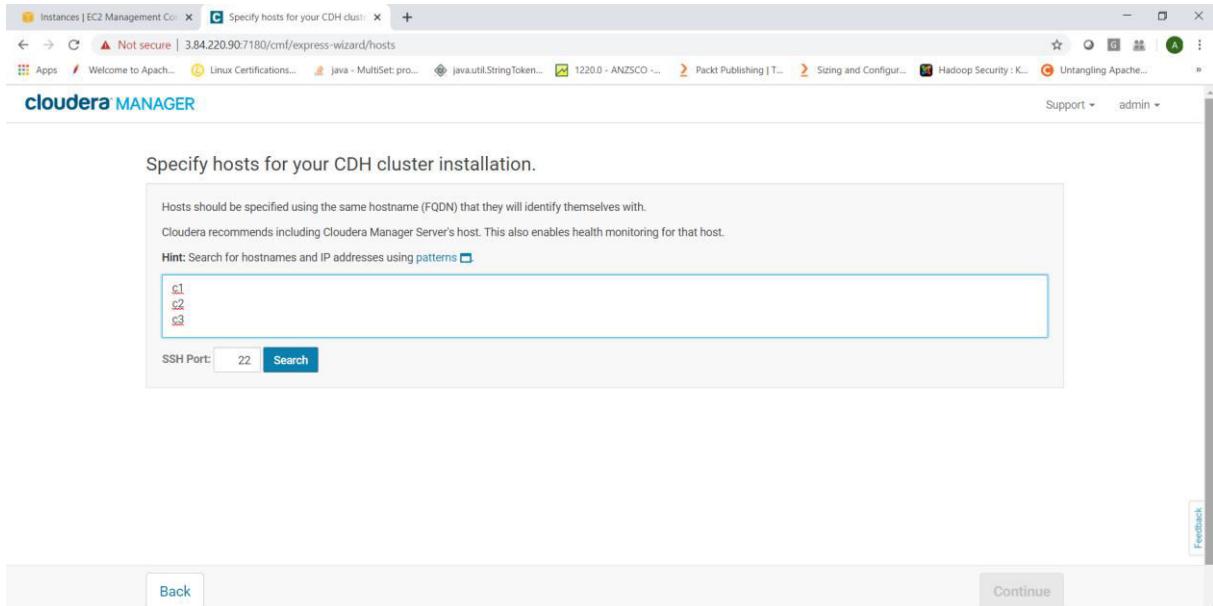
Access License Agreement and click on continue

By default “Cloudera Enterprise Trial” edition would be selected, click on continue



Click on continue

And provide hostnames as updated in /etc/hosts



Click on search

Specify hosts for your CDH cluster installation.

Hosts should be specified using the same hostname (FQDN) that they will identify themselves with. Cloudera recommends including Cloudera Manager Server's host. This also enables health monitoring for that host.

Hint: Search for hostnames and IP addresses using patterns

3 hosts scanned, 3 running SSH. [New Search](#)

Hostname (FQDN)	IP Address	Currently Managed	Result
c1	172.31.95.168	No	Host ready: 14 ms response time.
c2	172.31.90.165	No	Host ready: 1 ms response time.
c3	172.31.82.22	No	Host ready: 1 ms response time.

[Back](#) [Continue](#)

Click on continue

Here let the selection be parcels as recommended and scroll down to see options available for other parcels, although we can later add any parcels, so as of now click on continue.

Cluster Installation

Select Repository

Cloudera recommends the use of parcels for installation over packages, because parcels enable Cloudera Manager to easily manage the software on your cluster, automating the deployment and upgrade of service binaries. Electing not to use parcels will require you to manually upgrade packages on all hosts in your cluster when software updates are available, and will prevent you from using Cloudera Manager's rolling upgrade capabilities.

Choose Method Use Packages [?](#) Use Parcels (Recommended) [?](#) [More Options](#) [Proxy Settings](#)

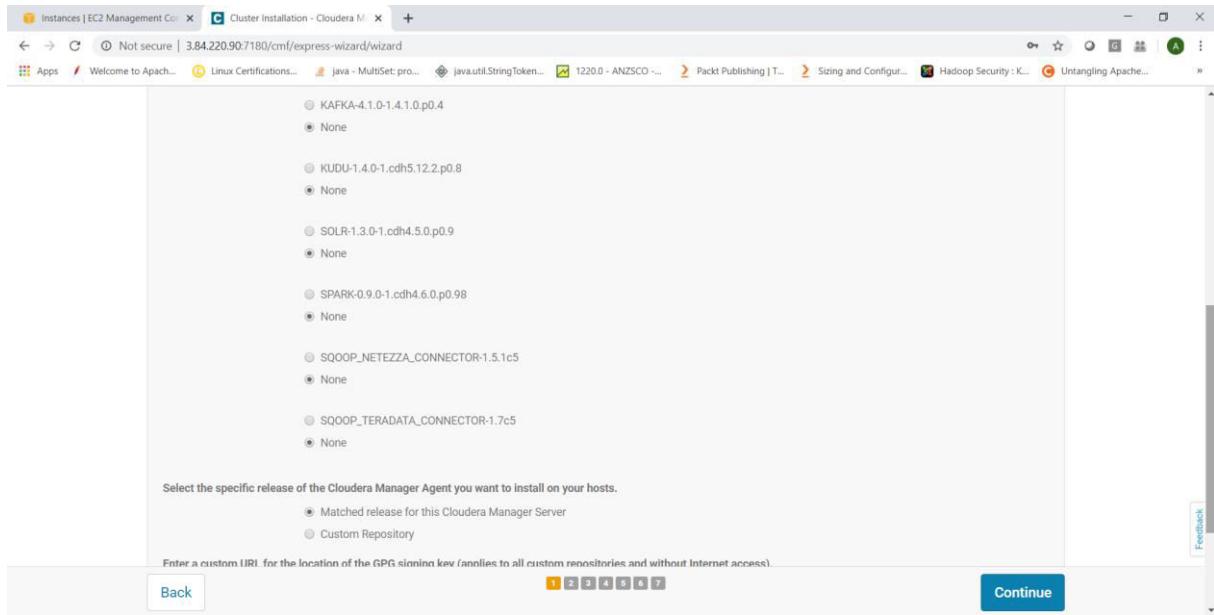
CDH Version CDH-5.16.2-1.cdh5.16.2.p0.8 CDH-4.7.1-1.cdh4.7.1.p0.47

Versions of CDH that are too new for this version of Cloudera Manager (5.16.2) will not be shown.

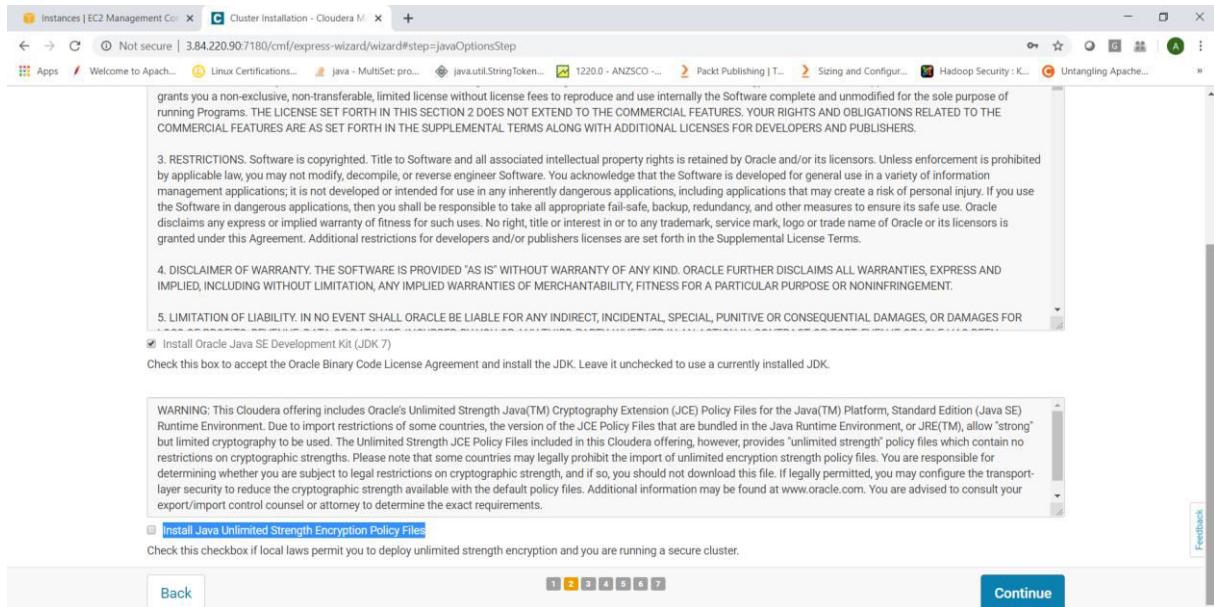
Additional Parcels ACCUMULO-1.7.2-5.5.0.ACCUMUL05.5.0.p0.8 ACCUMULO-1.4.4-1.cdh4.5.0.p0.65 None

IMPALA-2.1.0-1.impala2.0.0.p0.1995 None

[Back](#) [Continue](#)



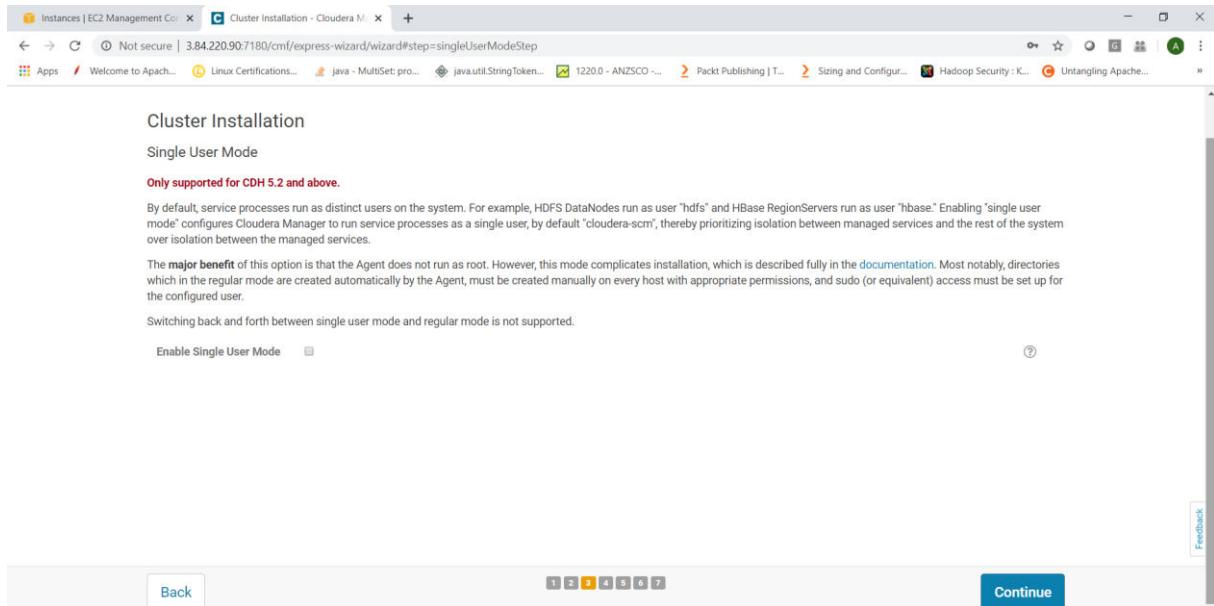
Accept JDK license



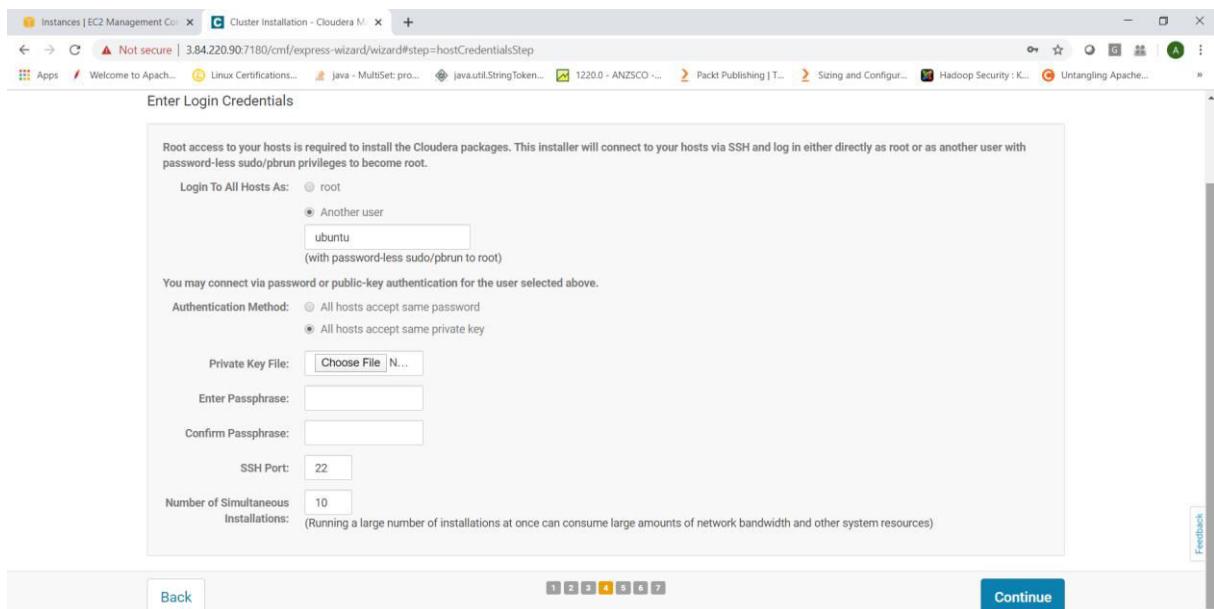
You can optionally select Install Java Unlimited Strength Encryption Policy Files, however, that can be chosen later when needed.

Click on continue

On single user mode, leave it as default and click on continue



Here we will now provide our user and the “*.pem” file which we have downloaded when we created key pair.



Default user for ubuntu machines is “ubuntu” and provide your “*.pem” file.

Note** Irrespective of whatever user & passphrase you created when converting pem to ppk, use default user for ec2 instance and pem as private key to connect.

Click on continue

Instances | EC2 Management Consoles Cluster Installation - Cloudera Manager

Not secure | 3.84.220.90:7180/cmfeexpress-wizard/wizard#step=installStep

Apps / Welcome to Apache... / Linux Certifications... / Java - MultiSet: pro... / java.util.StringToken... / 1220.0 - ANZSCO ... / Packt Publishing | T... / Sizing and Configur... / Hadoop Security : K... / Untangling Apache...

Cluster Installation

Install Agents

Installation in progress.

0 of 3 host(s) completed successfully. [Abort Installation](#)

Hostname	IP Address	Progress	Status	Details
c1	172.31.95.168	<div style="width: 100%;"> </div>	Refreshing package metadata...	Details
c2	172.31.90.165	<div style="width: 100%;"> </div>	Refreshing package metadata...	Details
c3	172.31.82.22	<div style="width: 100%;"> </div>	Refreshing package metadata...	Details

Feedback

Back Continue

1 2 3 4 5 6 7

Instances | EC2 Management Consoles Cluster Installation - Cloudera Manager

Not secure | 3.84.220.90:7180/cmfeexpress-wizard/wizard#step=installStep

Apps / Welcome to Apache... / Linux Certifications... / Java - MultiSet: pro... / java.util.StringToken... / 1220.0 - ANZSCO ... / Packt Publishing | T... / Sizing and Configur... / Hadoop Security : K... / Untangling Apache...

Cluster Installation

Install Agents

Installation in progress.

0 of 3 host(s) completed successfully. [Abort Installation](#)

Hostname	IP Address	Progress	Status	Details
c1	172.31.95.168	<div style="width: 100%;"> </div>	Installing cloudera-manager-agent package...	Details
c2	172.31.90.165	<div style="width: 100%;"> </div>	Installing oracle-j2sdk1.7 package...	Details
c3	172.31.82.22	<div style="width: 100%;"> </div>	Installing oracle-j2sdk1.7 package...	Details

Feedback

Back Continue

1 2 3 4 5 6 7

Instances | EC2 Management Consoles Cluster Installation - Cloudera Manager

Not secure | 3.84.220.90:7180/cmfeexpress-wizard/wizard#step=installStep

Apps / Welcome to Apache... / Linux Certifications... / Java - MultiSet: pro... / java.util.StringToken... / 1220.0 - ANZSCO ... / Packt Publishing | T... / Sizing and Configur... / Hadoop Security : K... / Untangling Apache...

Cluster Installation

Install Agents

Installation in progress.

0 of 3 host(s) completed successfully. [Abort Installation](#)

Hostname	IP Address	Progress	Status	Details
c1	172.31.95.168	<div style="width: 100%;"> </div>	Installing cloudera-manager-agent package...	Details
c2	172.31.90.165	<div style="width: 100%;"> </div>	Installing cloudera-manager-agent package...	Details
c3	172.31.82.22	<div style="width: 100%;"> </div>	Installing cloudera-manager-agent package...	Details

Feedback

Back Continue

1 2 3 4 5 6 7

Cluster Installation

Install Agents

Installation in progress.

1 of 3 host(s) completed successfully. [Abort Installation](#)

Hostname	IP Address	Progress	Status	Details
c1	172.31.95.168	<div style="width: 100%;"> </div>	✓ Installation completed successfully.	Details
c2	172.31.90.165	<div style="width: 50%; background-color: #0070C0; color: white;"> </div>	○ Installing cloudera-manager-agent package...	Details
c3	172.31.82.22	<div style="width: 50%; background-color: #0070C0; color: white;"> </div>	○ Installing cloudera-manager-agent package...	Details

Back Continue

Cluster Installation

Install Agents

Installation completed successfully.

3 of 3 host(s) completed successfully.

Hostname	IP Address	Progress	Status	Details
c1	172.31.95.168	<div style="width: 100%;"> </div>	✓ Installation completed successfully.	Details
c2	172.31.90.165	<div style="width: 100%;"> </div>	✓ Installation completed successfully.	Details
c3	172.31.82.22	<div style="width: 100%;"> </div>	✓ Installation completed successfully.	Details

Back Continue

Click on continue

Cluster Installation

Install Parcels

The selected parcels are being downloaded and installed on all the hosts in the cluster.

CDH 5.16.2-1.cdh5.16.2.p0.8	Downloaded: 33%	Distributed: 0/0	Unpacked: 0/0	Activated: 0/0
-----------------------------	-----------------	------------------	---------------	----------------

Wait for parcels to be downloaded + distributed + unpacked and activated

If you have any connectivity issues, wait and see if those resolved.

In previous step, cloudera-scm-agent would be installed in all your instances and that would be already sending heartbeats to cloudera-scm-server running on your first instance (i.e. c1)

You can optionally connect to instance and check for status of “cloudera-scm-agent” on each instance.

```
root@ip-172-31-95-168:~#
root@ip-172-31-95-168:~# vi /etc/hosts
root@ip-172-31-95-168:~# vi /etc/hosts
root@ip-172-31-95-168:~# wget https://archive.cloudera.com/cm5/installer/5.16/cloudera-manager-installer.bin
--2020-05-04 01:06:14 -- https://archive.cloudera.com/cm5/installer/5.16/cloudera-manager-installer.bin
Resolving archive.cloudera.com (archive.cloudera.com) ... 151.101.248.167
Connecting to archive.cloudera.com (archive.cloudera.com)|151.101.248.167|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 521501 (509K) [application/octet-stream]
Saving to: 'cloudera-manager-installer.bin'

cloudera-manager-installer.bin 100%[=====] 509.28K ---KB/s   in 0.01s
2020-05-04 01:06:14 (48.8 MB/s) - 'cloudera-manager-installer.bin' saved [521501/521501]

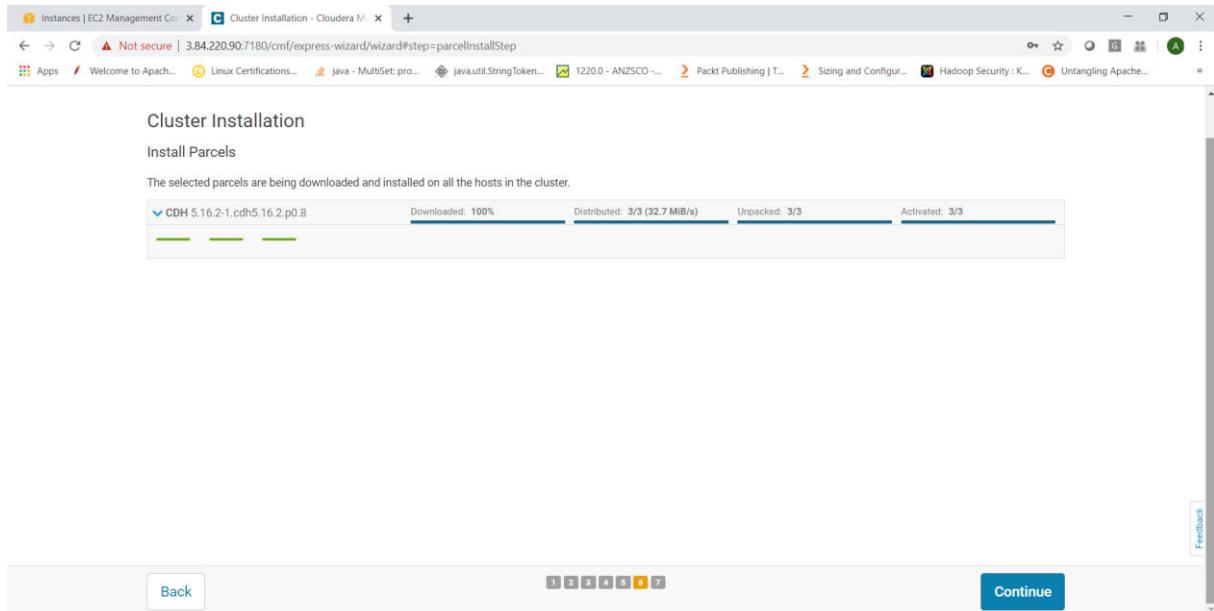
root@ip-172-31-95-168:~# chmod -R 777 cloudera-manager-installer.bin
root@ip-172-31-95-168:~# ls
cloudera-manager-installer.bin  snap
root@ip-172-31-95-168:~# ./cloudera-manager-installer.bin
root@ip-172-31-95-168:~# service cloudera-scm-server status
● cloudera-scm-server.service - LSB: Cloudera SCM Server
  Loaded: loaded (/etc/init.d/cloudera-scm-server; bad; vendor preset: enabled)
  Active: active (exited) since Mon 2020-05-04 01:09:48 UTC; 1min 14s ago
    Docs: man:systemd-sysv-generator(8)
  Process: 14751 ExecStart=/etc/init.d/cloudera-scm-server start (code=exited, status=0/SUCCESS)

May 04 01:09:43 ip-172-31-95-168 systemd[1]: Starting LSB: Cloudera SCM Server...
May 04 01:09:43 ip-172-31-95-168 su[14773]: Successful su for cloudera-scm by root
May 04 01:09:43 ip-172-31-95-168 su[14773]: + ??? root:cloudera-scm
May 04 01:09:43 ip-172-31-95-168 su[14773]: pam_unix(su:session): session opened for user cloudera-scm by (uid=0)
May 04 01:09:48 ip-172-31-95-168 cloudera-scm-server[14751]: Starting cloudera-scm-server: * cloudera-scm-server started
May 04 01:09:48 ip-172-31-95-168 systemd[1]: Started LSB: Cloudera SCM Server.
root@ip-172-31-95-168:~# service cloudera-scm-agent status
● cloudera-scm-agent.service - LSB: Cloudera SCM Agent
  Loaded: loaded (/etc/init.d/cloudera-scm-agent; bad; vendor preset: enabled)
  Active: active (exited) since Mon 2020-05-04 01:23:00 UTC; 4min 14s ago
    Docs: man:systemd-sysv-generator(8)
  Process: 23978 ExecStart=/etc/init.d/cloudera-scm-agent start (code=exited, status=0/SUCCESS)

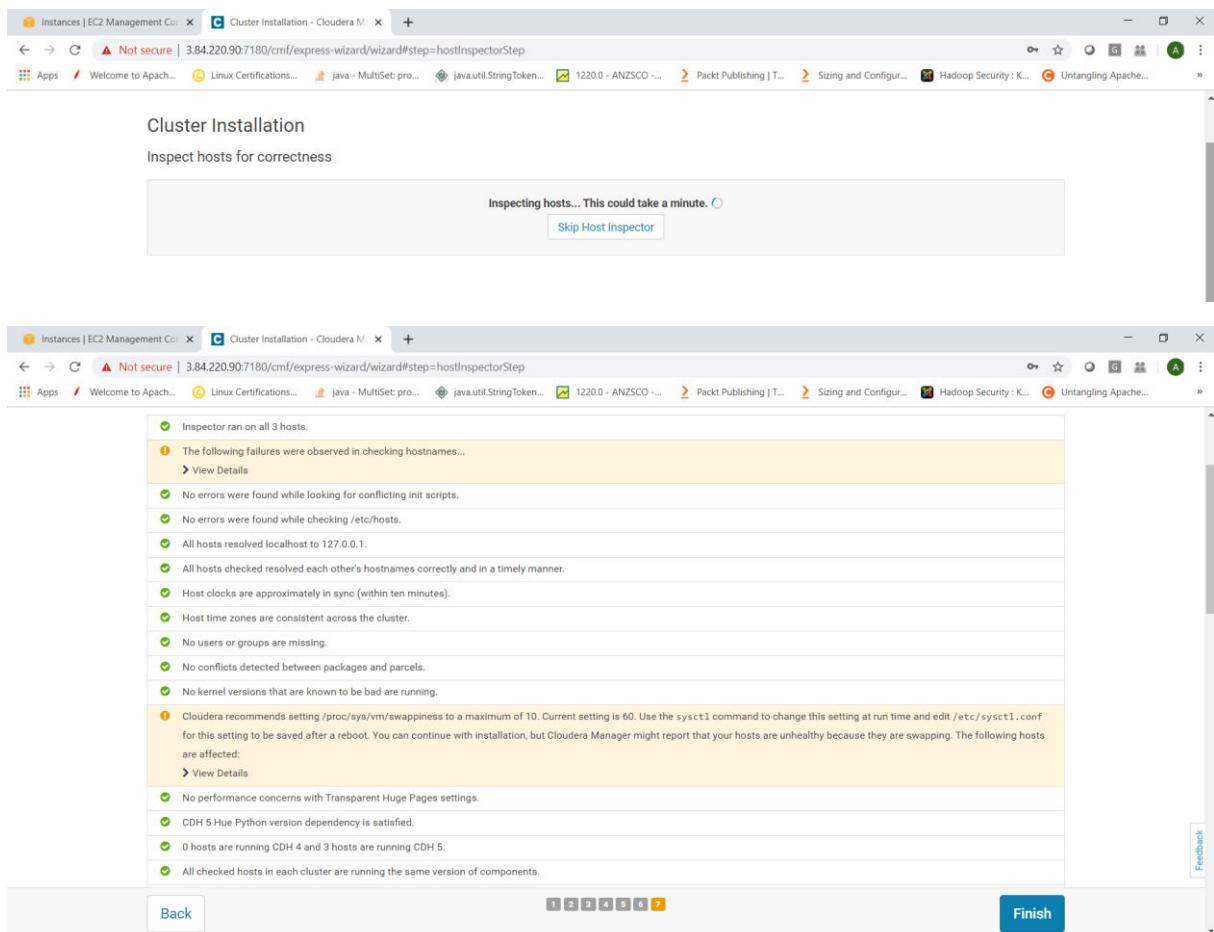
May 04 01:22:59 ip-172-31-95-168 systemd[1]: Starting LSB: Cloudera SCM Agent...
May 04 01:22:59 ip-172-31-95-168 su[23998]: Successful su for root by root
May 04 01:22:59 ip-172-31-95-168 su[23998]: + ??? root:root
May 04 01:22:59 ip-172-31-95-168 su[23998]: pam_unix(su:session): session opened for user root by (uid=0)
May 04 01:23:00 ip-172-31-95-168 cloudera-scm-agent[23978]: Starting cloudera-scm-agent: * cloudera-scm-agent started
May 04 01:23:00 ip-172-31-95-168 systemd[1]: Started LSB: Cloudera SCM Agent.
root@ip-172-31-95-168:~#
```

Back to admin console





Click on continue



Let's set swappiness on instances to 10 and then run host inspector again

```
May 04 01:22:59 ip-172-31-95-168 su[23990]: pam_unix(su:session): session opened for user root by (uid=0)
May 04 01:23:00 ip-172-31-95-168 cloudera-scm-agent[23978]: Starting cloudera-scm-agent: * cloudera-scm-agent started
May 04 01:23:00 ip-172-31-95-168 systemd[1]: Started LSB: Cloudera SCM Agent.
root@ip-172-31-95-168:~# vi /etc/sysctl.conf
```

```
root@ip-172-31-95-168:~  
# /etc/sysctl.conf - Configuration file for setting system variables  
# See /etc/sysctl.d/ for additional system variables.  
# See sysctl.conf(5) for information.  
  
vm.swappiness = 10  
kernel.concurrence = acceptconn  
  
# Recommend the following to stop low-level messages on interface  
kernel.sysrq=3  
  
*****  
# Functions previously found in netbase  
  
# recommend the next two lines to enable kernel parameter (reverse-path filter)  
# from or Source Address Verification to sit interfaces to  
# prevent most spoofing attempts  
net.ipv4.conf.default.rp_filter=1  
net.ipv4.conf.all.rp_filter=1  
  
# recommend the next line to enable TCP/IP SYN cookies  
# See https://www.netfilter.org/wiki/Ruleset/  
# Note this may impact TCP SYN sending too  
net.ipv4.tcp_syncookies=1  
  
# recommend the next line to enable packet forwarding for ipvs  
net.ipv4.ip_forward=1  
  
# recommend the next line to enable packet forwarding for ipvn  
# Reboot after this change disables Stateless Address Autoconfiguration  
# based on Router Advertisements for this host  
net.ipv6.conf.all.forwarding=1  
  
*****  
# Additional settings - These settings can improve the network  
# security of the host and protect against some network attacks  
# including spoofing attacks and MITM in the middle attacks through  
# disconnection. Some network environments, however, require that these  
# settings are disabled or review and enable them as needed.  
  
# Do not accept ICMP redirects (protect MITM attack)  
net.ipv4.icmp_echo_ignore_all=1  
net.ipv4.icmp_echo_ignore_broadcasts=1  
-- INSERT --
```

:wq

```
root@ip-172-31-95-168:~# vi /etc/sysctl.conf
root@ip-172-31-95-168:~# sysctl -p
Vm.swappiness = 10
root@ip-172-31-95-168:~#
```

Similarly on all instances

Let's also check if NTP service is installed on each instance and running, if not install and start it (to ensure all machines are in sync)

```
root@ip-172-31-95-168:~# apt-get install ntp
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following additional packages will be installed:
  libopts25
Suggested packages:
  ntp-doc
The following NEW packages will be installed:
  libopts25 0p
0 upgraded, 2 newly installed, 0 to remove and 11 not upgraded.
Need to get 575 kB of archives.
After this operation, 1,792 kB of additional disk space will be used.
Do you want to continue? [Y/n] Y
Get:1 http://us-east-1.ec2.archive.ubuntu.com/ubuntu xenial/main amd64 libopts25 amd64 1:5.18.7-3 [57.8 kB]
Get:2 http://us-east-1.ec2.archive.ubuntu.com/ubuntu xenial-updates/main amd64 ntp amd64 1:4.2.8p4+dfsg-3ubuntu5.10 [517 kB]
Fetched 575 kB in 0s (4,087 kB/s)
Selecting previously unselected package libopts25:amd64.
(Reading database ... 73032 files and directories currently installed.)
Preparing to unpack .../libopts25_1%3a5.18.7-3_amd64.deb ...
Unpacking libopts25:amd64 (1:5.18.7-3) ...
Selecting previously unselected package ntp.
Preparing to unpack .../ntp_1%3a4.2.8p4+dfsg-3ubuntu5.10_amd64.deb ...
Unpacking ntp (1:4.2.8p4+dfsg-3ubuntu5.10) ...
Processing triggers for libc-bin (2.23-0ubuntu11) ...
Processing triggers for man-db (2.7.5-1) ...
Processing triggers for ureadahead (0.100.0-19.1) ...
Processing triggers for systemd (229-4ubuntu21.27) ...
Setting up libopts25:amd64 (1:5.18.7-3) ...
Setting up ntp (1:4.2.8p4+dfsg-3ubuntu5.10) ...
Processing triggers for libc-bin (2.23-0ubuntu11) ...
Processing triggers for ureadahead (0.100.0-19.1) ...
Processing triggers for systemd (229-4ubuntu21.27) ...
root@ip-172-31-95-168:~# █

root@ip-172-31-95-168:~# service ntp start
root@ip-172-31-95-168:~# service ntp status
● ntp.service - LSB: Start NTP daemon
  Loaded: loaded (/etc/init.d/ntp; bad; vendor preset: enabled)
  Active: active (running) since Mon 2020-05-04 01:35:43 UTC; 31s ago
    Docs: man:systemd-sysv-generator(8)
   CGroup: /system.slice/ntp.service
           └─26613 /usr/sbin/ntpd -p /var/run/ntpd.pid -g -u 114:119

May 04 01:35:43 ip-172-31-95-168 ntpd[26613]: Listen normally on 3 eth0 172.31.95.168:123
May 04 01:35:43 ip-172-31-95-168 ntpd[26613]: Listen normally on 4 lo ::1:123
May 04 01:35:43 ip-172-31-95-168 ntpd[26613]: Listen normally on 5 eth0 [fe80::1072:d8ff:fe3:3b23%2]:123
May 04 01:35:43 ip-172-31-95-168 ntpd[26613]: Listening on routing socket on fd #22 for interface updates
May 04 01:35:44 ip-172-31-95-168 ntpd[26613]: Soliciting pool server 136.243.7.20
May 04 01:35:45 ip-172-31-95-168 ntpd[26613]: Soliciting pool server 204.11.201.12
May 04 01:35:46 ip-172-31-95-168 ntpd[26613]: Soliciting pool server 198.60.22.240
May 04 01:35:47 ip-172-31-95-168 ntpd[26613]: Soliciting pool server 216.126.233.109
May 04 01:35:48 ip-172-31-95-168 ntpd[26613]: Soliciting pool server 91.189.89.199
May 04 01:36:11 ip-172-31-95-168 systemd[1]: Started LSB: Start NTP daemon.
root@ip-172-31-95-168:~# █
```

Note** if using centos/redhat machines, you will have to disable SELINUX in /etc/sysconfig. Since I am using Ubuntu, its not required.

Now back to admin console and rerun “host inspector”

Host Inspector Results	
The following failures were observed in checking hostnames...	
View Details	
<ul style="list-style-type: none"> No errors were found while looking for conflicting init scripts. No errors were found while checking /etc/hosts. All hosts resolved localhost to 127.0.0.1. All hosts checked resolved each other's hostnames correctly and in a timely manner. Host clocks are approximately in sync (within ten minutes). Host time zones are consistent across the cluster. No users or groups are missing. No conflicts detected between packages and parcels. No kernel versions that are known to be bad are running. No problems were found with /proc/sys/vm/swappiness on any of the hosts. No performance concerns with Transparent Huge Pages settings. CDH 5 Hue Python version dependency is satisfied. 0 hosts are running CDH 4 and 3 hosts are running CDH 5. All checked hosts in each cluster are running the same version of components. All managed hosts have consistent versions of Java. All checked Cloudera Management Daemons versions are consistent with the server. All checked Cloudera Management Agents versions are consistent with the server. 	
Back	Finish

Version Summary					
Cluster 1					
All Hosts					
c1[3]	Component	Version	Hosts	Release	CDH Version
	Supervisor	3.0-cm5.16.2	All Hosts	Unavailable	Not applicable
	Bigtop-Tomcat (CDH 5 only)	0.7.0+cdh5.16.2+0	All Hosts	1.cdh5.16.2.p0.22	CDH 5
	Cloudera Manager Agent	5.16.2	All Hosts	1.cm5162.p0.7	Not applicable
	Cloudera Manager	5.16.2	All Hosts	1.cm5162.p0.7	Not applicable
	Management Daemons				
	Crunch (CDH 5 only)	0.11.0+cdh5.16.2+106	All Hosts	1.cdh5.16.2.p0.23	CDH 5
	Flume NG	1.6.0+cdh5.16.2+196	All Hosts	1.cdh5.16.2.p0.23	CDH 5
	Hadoop	2.6.0+cdh5.16.2+2863	All Hosts	1.cdh5.16.2.p0.26	CDH 5
	MapReduce 1	2.6.0+cdh5.16.2+2863	All Hosts	1.cdh5.16.2.p0.26	CDH 5
	HDFS	2.6.0+cdh5.16.2+2863	All Hosts	1.cdh5.16.2.p0.26	CDH 5
	HttpFS	2.6.0+cdh5.16.2+2863	All Hosts	1.cdh5.16.2.p0.26	CDH 5
	hadoop-kms	2.6.0+cdh5.16.2+2863	All Hosts	1.cdh5.16.2.p0.26	CDH 5
	MapReduce 2	2.6.0+cdh5.16.2+2863	All Hosts	1.cdh5.16.2.p0.26	CDH 5
	YARN	2.6.0+cdh5.16.2+2863	All Hosts	1.cdh5.16.2.p0.26	CDH 5

Check versions of software packages and click on Finish

The screenshot shows the 'Cluster Setup' section of the Cloudera Manager interface. Under 'Select Services', it lists various service combinations:

- Core Hadoop: HDFS, YARN (MapReduce 2 Included), ZooKeeper, Oozie, Hive, and Hue
- Core with HBase: HDFS, YARN (MapReduce 2 Included), ZooKeeper, Oozie, Hive, Hue, and HBase
- Core with Impala: HDFS, YARN (MapReduce 2 Included), ZooKeeper, Oozie, Hive, Hue, and Impala
- Core with Search: HDFS, YARN (MapReduce 2 Included), ZooKeeper, Oozie, Hive, Hue, and Solr
- Core with Spark: HDFS, YARN (MapReduce 2 Included), ZooKeeper, Oozie, Hive, Hue, and Spark
- All Services: HDFS, YARN (MapReduce 2 Included), ZooKeeper, Oozie, Hive, Hue, HBase, Impala, Solr, Spark, and Key-Value Store Indexer
- Custom Services: Choose your own services. Services required by chosen services will automatically be included. Flume can be added after your initial cluster has been set up.

At the bottom, there are 'Back' and 'Continue' buttons, along with a feedback link.

In cluster setup , let's choose Core Hadoop as of now. We can optionally choose any option or later add services to our cluster.

Click on Continue and it takes to “role assignment” page.

Assign roles and distribute roles across machines for your cluster services.

The screenshot shows the 'Assign Roles' section of the Cloudera Manager interface. It lists role assignments for different services:

- HDFS**:
 - NameNode x 1 New: c1
 - SecondaryNameNode x 1 New: c1
 - Balancer x 1 New: c1
 - HttpFS: Select hosts
 - NFS Gateway: Select hosts
 - DataNode x 2 New: c2[c3]
- Hive**:
 - Gateway x 3 New: c1[c2:c3]
 - Hive Metastore Server x 1 New: c1
 - WebHCat Server: Select hosts
 - HiveServer2 x 1 New: c1
- Hue**:
 - Hue Server x 1 New: c1
 - Load Balancer x 1 New: c1
- Cloudera Management Service**:
 - Service Monitor x 1 New: c1
 - Activity Monitor: Select a host
 - Host Monitor x 1 New: c1
 - Reports Manager x 1 New: c1
 - Event Server x 1 New: c1
 - Alert Publisher x 1 New: c1
 - Telemetry Publisher: Select a host
- Oozie**:
 - None

At the bottom, there are 'Back' and 'Continue' buttons, along with a feedback link.

For example , clicking in box for secondarynamenode shows

Cluster Setup

Assignment Roles

1 Host Selected

Select hosts for a new or existing role. The host list is filtered to remove hosts that are not valid candidates; these include hosts that are unhealthy, members of other clusters, or have an incompatible version of CDH installed on them.

Enter hostnames: host01, host[01-10], IP addresses or rack.

Hostname	IP Address	Rack	Cores	Physical Memory	Existing Roles	Added Roles
c1	172.31.95.168	/default	2	7.8 GB	<input type="checkbox"/> NN <input type="checkbox"/> B <input type="checkbox"/> G <input type="checkbox"/> HMS <input type="checkbox"/> HS <input type="checkbox"/> LB <input type="checkbox"/> SM <input type="checkbox"/> HM <input type="checkbox"/> RM	<input type="checkbox"/> ES <input type="checkbox"/> AP <input type="checkbox"/> OS <input type="checkbox"/> NM <input type="checkbox"/> SNN
c2	172.31.90.165	/default	2	7.8 GB	<input checked="" type="checkbox"/> DN <input type="checkbox"/> G <input type="checkbox"/> NM <input type="checkbox"/> SNN	
c3	172.31.82.22	/default	2	7.8 GB	<input type="checkbox"/> DN <input type="checkbox"/> G <input type="checkbox"/> NM	

Displaying 1 - 3 of 3

Cancel

This is how my role assignment looks

Instances | EC2 Management Consoles | Cluster Setup - Cloudera Manager | +

Not secure | 3.84.220.90:7180/cmfl/clusters/1/express-add-services/index#step=roleAssignmentsStep

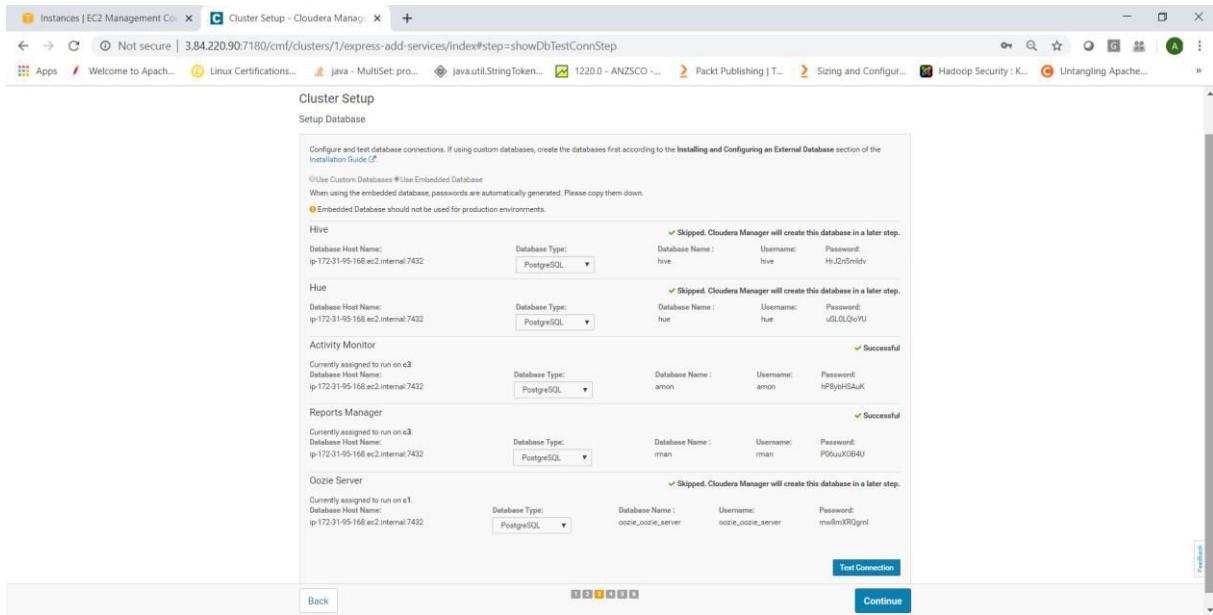
Apps Welcome to Apache... Linux Certifications... Java - MultiSet: pro... java.util.StringToken... 1220.0 - ANZSCO ... Packt Publishing | T... Sizing and Configur... Hadoop Security : K... Untangling Apache...

<input type="checkbox"/> NameNode x 1 New c1	<input type="checkbox"/> Secondary NameNode x 1 New c2+	<input type="checkbox"/> Balancer x 1 New c1	<input type="checkbox"/> HDFS x 3 New c1 2 3+
<input type="checkbox"/> NFS Gateway x 3 New c1 2 3	<input type="checkbox"/> DataNode x 3 New All Hosts		
<input type="checkbox"/> Hive Gateway x 3 New c1 2 3	<input type="checkbox"/> Hive Metastore Server x 1 New c2+	<input type="checkbox"/> WebHCat Server x 1 New c2+	<input type="checkbox"/> HiveServer2 x 1 New c2+
<input type="checkbox"/> Hue Hue Server x 1 New c1	<input type="checkbox"/> Load Balancer x 1 New c1		
<input type="checkbox"/> Cloudera Management Service Service Monitor x 1 New c3	<input type="checkbox"/> Activity Monitor x 1 New c3	<input type="checkbox"/> Host Monitor x 1 New c3	<input type="checkbox"/> Reports Manager x 1 New c2
<input type="checkbox"/> Event Server x 1 New c3	<input type="checkbox"/> Alert Publisher x 1 New c3	<input type="checkbox"/> Telemetry Publisher x 1 New c3	
<input type="checkbox"/> Oozie Oozie Server x 1 New c1			
<input type="checkbox"/> YARN (MR2 Included) ResourceManager x 1 New c3	<input type="checkbox"/> JobHistory Server x 1 New c3	<input type="checkbox"/> NodeManager x 3 New c1 2 3+	
<input type="checkbox"/> ZooKeeper Server x 3 New c1 2 3			

Back

Click on continue , this takes you to “Setup Database”. Here by default “postgres” is chosen. We can find these settings (username/passwds) from instances too.

Click on Test connection



Click on continue

Meanwhile connect to your first instance i.e. C1 (where cloudera-scm-server is running)

```
root@ip-172-31-95-168:~# cat /etc/cloudera-scm-server/db.mgmt.properties
# Auto-generated by initialize_embedded_db.sh
#
# 20200504-010938

# These are database credentials for databases
# created by "cloudera-scm-server-db" for
# Cloudera Manager Management Services,
# to be used during the installation wizard if
# the embedded database route is taken.

# The source of truth for these settings
# is the Cloudera Manager databases and
# changes made here will not be reflected
# there automatically.

com.cloudera.cmf.ACTIVITYMONITOR.db.type=postgresql
com.cloudera.cmf.ACTIVITYMONITOR.db.host=ip-172-31-95-168.ec2.internal:7432
com.cloudera.cmf.ACTIVITYMONITOR.db.name=amon
com.cloudera.cmf.ACTIVITYMONITOR.db.user=amon
com.cloudera.cmf.ACTIVITYMONITOR.db.password=hP8ybHSAuK
com.cloudera.cmf.REPORTSMANAGER.db.type=postgresql
com.cloudera.cmf.REPORTSMANAGER.db.host=ip-172-31-95-168.ec2.internal:7432
com.cloudera.cmf.REPORTSMANAGER.db.name=rman
com.cloudera.cmf.REPORTSMANAGER.db.user=rman
com.cloudera.cmf.REPORTSMANAGER.db.password=P06uuXOB4U
com.cloudera.cmf.NAVIGATOR.db.type=postgresql
com.cloudera.cmf.NAVIGATOR.db.host=ip-172-31-95-168.ec2.internal:7432
com.cloudera.cmf.NAVIGATOR.db.name=nav
com.cloudera.cmf.NAVIGATOR.db.user=nav
com.cloudera.cmf.NAVIGATOR.db.password=wOggPUOPN2
com.cloudera.cmf.NAVIGATORMETASERVER.db.type=postgresql
com.cloudera.cmf.NAVIGATORMETASERVER.db.host=ip-172-31-95-168.ec2.internal:7432
com.cloudera.cmf.NAVIGATORMETASERVER.db.name=navms
com.cloudera.cmf.NAVIGATORMETASERVER.db.user=navms
com.cloudera.cmf.NAVIGATORMETASERVER.db.password=pVKRTdZbIN
root@ip-172-31-95-168:~#
```

```
root@ip-172-31-95-168:~# cat /etc/cloudera-scm-server/db.properties
# Auto-generated by initialize_embedded_db.sh
#
# 20200504-010938

# These are database settings for CM Manager

com.cloudera.cmf.db.type=postgresql
com.cloudera.cmf.db.host=localhost:7432
com.cloudera.cmf.db.name=scm
com.cloudera.cmf.db.user=scm
com.cloudera.cmf.db.password=5BctQoRNva
com.cloudera.cmf.db.setupType=EMBEDDED
root@ip-172-31-95-168:~#
```

Back to admin console

Shows you default configurations for your services. You can review them or edit at this point of time.

As of now , let them be default and click on continue

Your cluster setup begins

The screenshot shows the 'Cluster Setup' page in the Cloudera Manager web interface. At the top, it says 'First Run Command'. Below that, the status is 'Running' and the date is 'May 4, 1:47:46 AM'. There is an 'Abort' button. A message indicates 'Completed 1 of 8 step(s)'. The steps listed are:

- > Ensuring that the expected software releases are installed on hosts.
- > Deploying Client Configuration
- > Start Cloudera Management Service, ZooKeeper
- > Start HDFS
- > Start YARN (MR2 Included)
- > Start Hive
- > Start Oozie
- > Start Hue

At the bottom, there are 'Back' and 'Continue' buttons, and a navigation bar with numbers 1 through 6.

This will configure settings, do the needful and start your services and thus your overall cluster.

Note** There might be various issues in cluster such as “heap settings”, “node offsets” etc and those issues can be resolved once cluster is up and running.

The screenshot shows the 'Cluster Setup' page in the Cloudera Manager web interface. The steps listed are:

- > Ensuring that the expected software releases are installed on hosts.
- > Deploying Client Configuration
- > Start Cloudera Management Service, ZooKeeper
- > Start HDFS
8/1 steps completed.
- > Start YARN (MR2 Included)
- > Start Hive
- > Start Oozie
- > Start Hue

The 'Start HDFS' step is expanded, showing its progress: 8/1 steps completed. The date for this step is 'May 4, 1:48:29 AM'. At the bottom, there are 'Back' and 'Continue' buttons, and a navigation bar with numbers 1 through 6.

You can expand any section to see what is happening

Instances | EC2 Management Consoles Cluster Setup - Cloudera Manager

Show All Steps Show Only Failed Steps Show Running Steps

> Ensuring that the expected software releases are installed on hosts.	May 4, 1:47:46 AM	53ms	
> Deploying Client Configuration	Cluster 1	May 4, 1:47:47 AM	15.63s
> Start Cloudera Management Service, ZooKeeper		May 4, 1:48:02 AM	26.52s
Start HDFS 0/1 steps completed.		May 4, 1:48:29 AM	
Execute 3 steps in sequence Waiting for command (Create /tmp Directory (73)) to finish		May 4, 1:48:29 AM	
> Formatting the name directories of the current NameNode. If the name directories are not empty, this is expected to fail.	NameNode (c1)	May 4, 1:48:29 AM	5.55s
> Start HDFS	HDFS	May 4, 1:48:34 AM	27.87s
> Creating HDFS /tmp directory Waiting for command (Create /tmp Directory (73)) to finish	NameNode (c1)	May 4, 1:49:02 AM	
> Start YARN (MR2 Included)			
> Start Hive			
> Start Oozie			
> Start Hue			

Feedback

Back Continue

Instances | EC2 Management Consoles Cluster Setup - Cloudera Manager

Show All Steps Show Only Failed Steps Show Running Steps

> Ensuring that the expected software releases are installed on hosts.	May 4, 1:47:46 AM	53ms	
> Deploying Client Configuration	Cluster 1	May 4, 1:47:47 AM	15.63s
> Start Cloudera Management Service, ZooKeeper		May 4, 1:48:02 AM	26.52s
Start HDFS Successfully completed 1 steps.		May 4, 1:48:29 AM	56.73s
Execute 3 steps in sequence Command (Create /tmp Directory (73)) has completed successfully		May 4, 1:48:29 AM	56.73s
> Formatting the name directories of the current NameNode. If the name directories are not empty, this is expected to fail.	NameNode (c1)	May 4, 1:48:29 AM	5.55s
> Start HDFS	HDFS	May 4, 1:48:34 AM	27.87s
> Creating HDFS /tmp directory	NameNode (c1)	May 4, 1:49:02 AM	23.27s
> Start YARN (MR2 Included)		May 4, 1:49:26 AM	28.51s
> Start Hive 0/1 steps completed.		May 4, 1:49:54 AM	
> Start Oozie			
> Start Hue			

Feedback

Back Continue

Instances | EC2 Management Consoles Cluster Setup - Cloudera Manager

Show All Steps Show Only Failed Steps Show Running Steps

First Run Command			
Status Finished May 4, 1:47:46 AM 4.5m			
Finished First Run of the following services successfully: ZooKeeper, HDFS, YARN (MR2 Included), Hive, Oozie, Hue, Cloudera Management Service.			
Completed 8 of 8 step(s).			
> Ensuring that the expected software releases are installed on hosts.	May 4, 1:47:46 AM	53ms	
> Deploying Client Configuration	Cluster 1	May 4, 1:47:47 AM	15.63s
> Start Cloudera Management Service, ZooKeeper		May 4, 1:48:02 AM	26.52s
> Start HDFS		May 4, 1:48:29 AM	56.73s
> Start YARN (MR2 Included)		May 4, 1:49:26 AM	28.51s
> Start Hive		May 4, 1:49:54 AM	55.57s
> Start Oozie		May 4, 1:50:50 AM	59.62s
Start Hue Successfully completed 1 steps.		May 4, 1:51:49 AM	27.34s
Execute 2 steps in sequence		May 4, 1:51:49 AM	27.34s

Feedback

Back Continue

The image shows two screenshots of the Cloudera Manager interface.

The top screenshot is titled "Cluster Setup" and displays a "Congratulations!" message: "The services are installed, configured, and running on your cluster." It includes a "Feedback" link in the bottom right corner.

The bottom screenshot is titled "Home - Cloudera Manager" and shows the main dashboard for "Cluster 1". It includes sections for "Clusters", "Hosts", "Diagnostics", "Audits", "Charts", "Backup", and "Administration". The "Clusters" section lists "3 Hosts", "HDFS", "Hive", "Hue", "Oozie", "YARN", and "ZooKeeper". The "Charts" section features two line graphs: "HDFS IO" and "Cluster Network IO". The "HDFS IO" graph shows rates of 1b/s, 2b/s, and 3b/s. The "Cluster Network IO" graph shows rates of 9.5M/s and 19.1M/s. A yellow banner at the top states: "You are running Cloudera Manager in non-production mode, which uses an embedded PostgreSQL database. Switch to using a supported external database before moving into production. More Details".

Your cluster is setup. Now we can resolve issues seen one by one.

