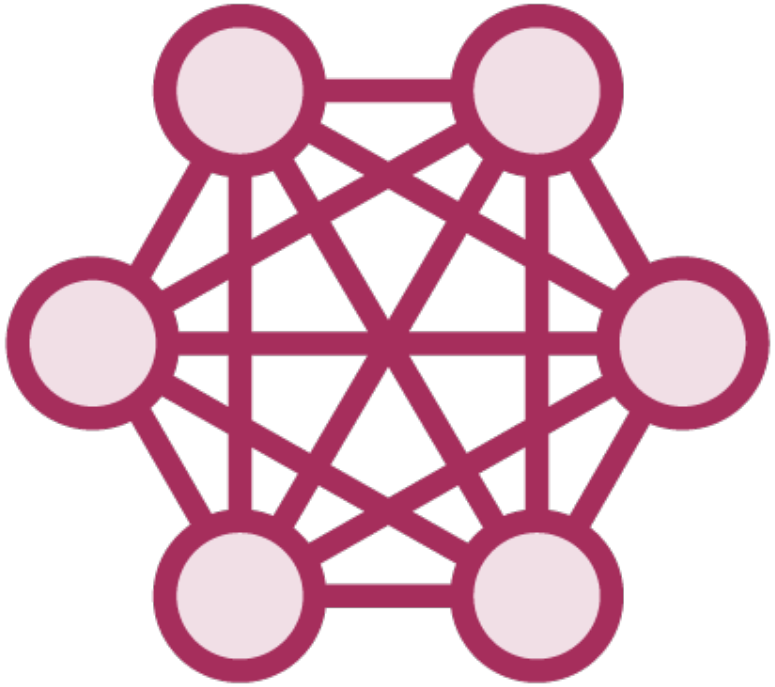
Prerequisite: Getting Linux Machines Ready for Your Cluster

Xavier Morera
PASSIONATE ABOUT TEACHING

@xmorera www.xaviermorera.com

Getting Ready for Big Data



We need our cluster

Define:cluster

- Set of connected machines that work together as a single system

Create a cluster

- Ready for Cloudera & Hadoop install

Did we mention OS?

Linux as OS of choice

#1 choice for Hadoop & many Open Source

Got Linux?

Get a base machine ready

- Fresh Linux install
- Optional snapshot or export appliance

Initial machine configuration

Network configuration

- NAT Network in this case

Clone machine to create your cluster

- Hostname, network configuration, ...

Now you have a base cluster!

CentOS

Community Enterprise Operating System

Free Linux distribution

RHEL compatible

- Red Hat Enterprise Linux

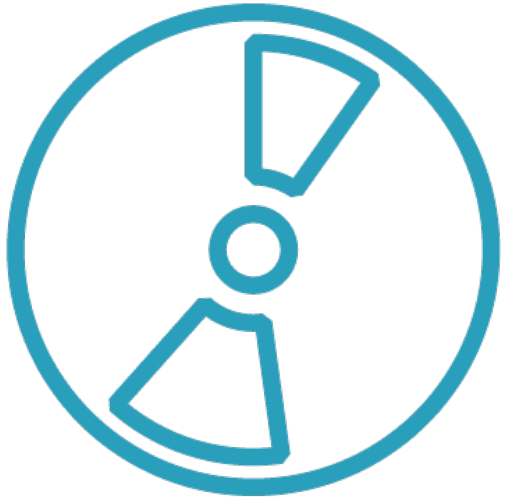
Supported Versions by Cloudera

- Yes: 7.1 & 7.2 plus some 6.x and 5.x
- No: 7.0

Let's get CentOS installed

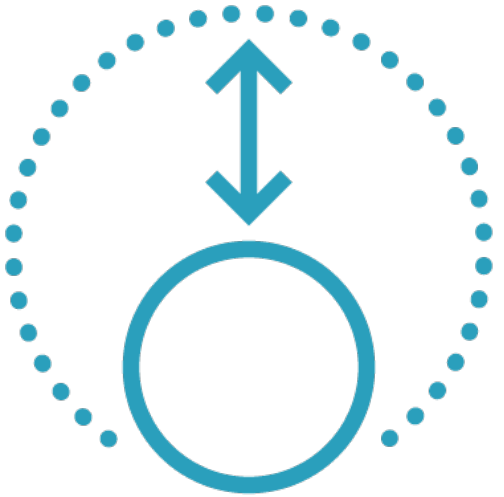
- From www.centos.org

Get CentOS: Which ISO?



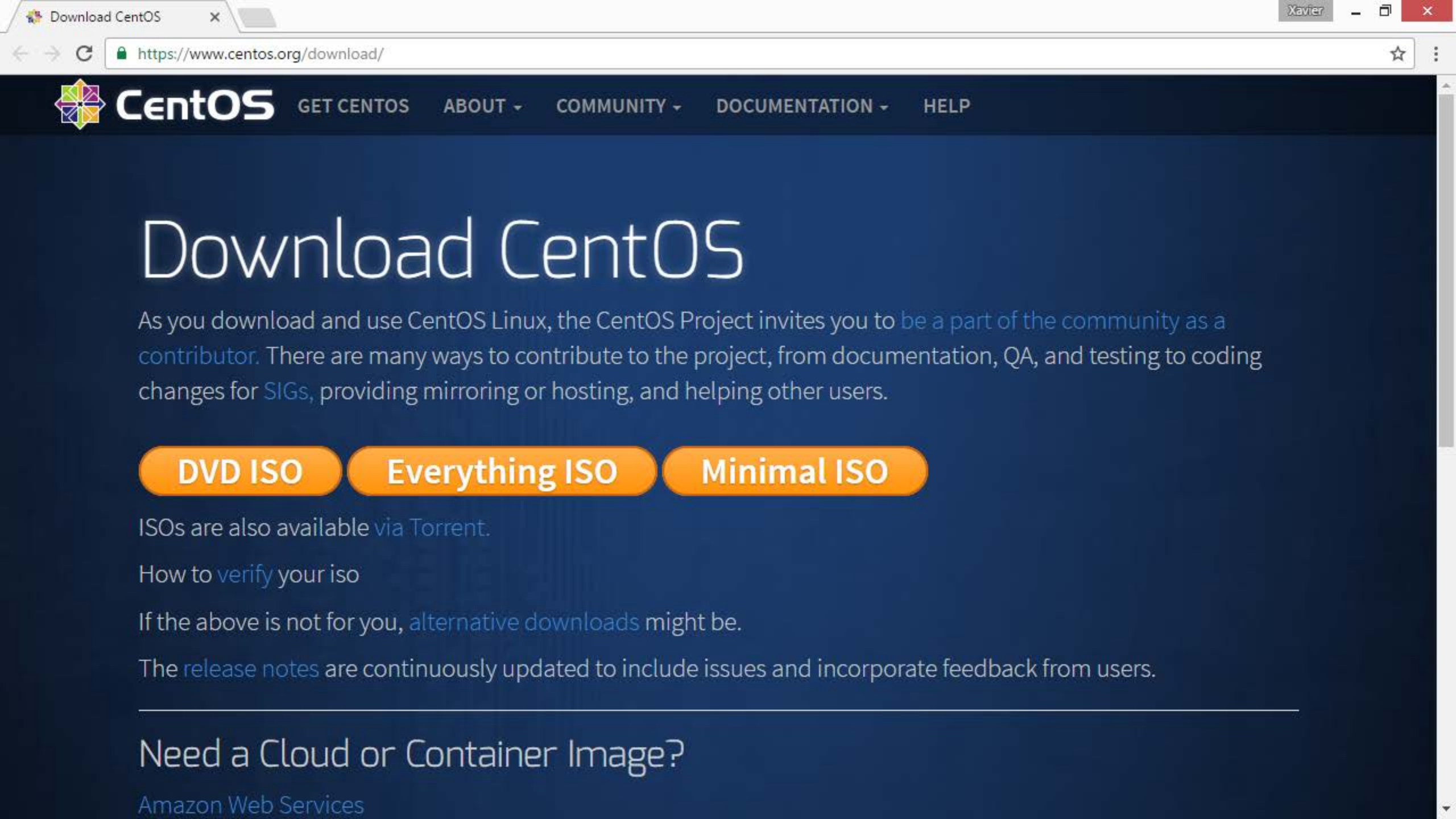
DVD ISO

Everything ISO



Minimal ISO

Start Downloading CentOS...



Download CentOS

https://www.centos.org/download/

CentOS GET CENTOS ABOUT COMMUNITY DOCUMENTATION HELP

Download CentOS

As you download and use CentOS Linux, the CentOS Project invites you to be a part of the community as a contributor. There are many ways to contribute to the project, from documentation, QA, and testing to coding changes for SIGs, providing mirroring or hosting, and helping other users.

DVD ISO | Everything ISO | Minimal ISO

ISOs are also available via Torrent.

How to verify your iso

If the above is not for you, alternative downloads might be.

The release notes are continuously updated to include issues and incorporate feedback from users.

Need a Cloud or Container Image?

Amazon Web Services

CentOS Mirror

isoredirect.centos.org/centos/7/isos/x86_64/CentOS-7-x86_64-DVD-1511.iso

CentOS

CentOS on the Web: Cen

In order to conserve the limited bandw

The following mirrors should have the

Actual Country -

http://mirrors.ucr.ac.cr/centos/7/isos/x86

Nearby Countries -

http://centos.brisanet.com.br/7/isos/x86_
http://mirror.ufscar.br/centos/7/isos/x86_
http://mirror.facom.ufms.br/centos/7/isos
http://centos.ufes.br/7/isos/x86_64/Cent
http://mirror.globo.com/centos/7/isos/x86
http://mirror.nbtelecom.com.br/centos/7/
http://mirror.gtdinternet.com/7/isos/x86_64/CentOS-7-x86_64-DVD-1511.iso
http://mirror.orbyta.com/7/isos/x86_64/CentOS-7-x86_64-DVD-1511.iso
ftp://ftp.inf.utfsm.cl/pub/Linux/CentOS/7/isos/x86_64/CentOS-7-x86_64-DVD-1511.iso
http://centos.ar.host-engine.com/7/isos/x86_64/CentOS-7-x86_64-DVD-1511.iso
http://centos.xfree.com.ar/7/isos/x86_64/CentOS-7-x86_64-DVD-1511.iso
http://mirrors.uprm.edu/centos/7/isos/x86_64/CentOS-7-x86_64-DVD-1511.iso
http://mirrors.mit.edu/centos/7/isos/x86_64/CentOS-7-x86_64-DVD-1511.iso
http://centos.den.host-engine.com/7/isos/x86_64/CentOS-7-x86_64-DVD-1511.iso
http://mirrors.lga7.us.voxel.net/centos/7/isos/x86_64/CentOS-7-x86_64-DVD-1511.iso
http://mirror.linuxfix.com/centos/7/isos/x86_64/CentOS-7-x86_64-DVD-1511.iso
http://centos.mirrors.tds.net/pub/linux/centos/7/isos/x86_64/CentOS-7-x86_64-DVD-1511.iso

Xavier

Downloads

Disc Image Tools

File | Home | Share | View | Manage

Navigation pane | Preview pane | Details pane | Panes

Extra large icons | Large icons | Medium icons | Small icons | List | Details | Layout

Sort by | Current view

Item check boxes | File name extensions | Hidden items | Hide selected items | Options | Show/hide

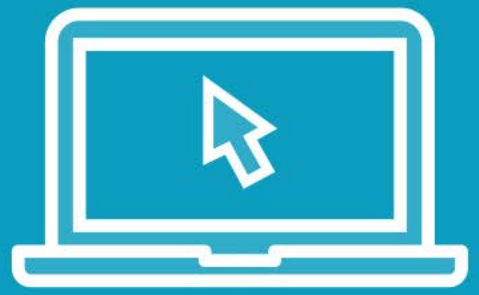
This PC ▸ Downloads

Search Downloads

VirtualBox-5.1.8-111374-Win.exe
cloudera-quickstart-vm-5.8.0-0-virtualbox.zip
archive
CentOS-7-x86_64-DVD-1511.iso

4 items | 1 item selected 4.03 GB

Demo



Exercise 5
Installing Linux from Scratch with CentOS

Oracle VM VirtualBox Manager

File Machine Help

New Settings Discard Start

Details Snapshots

cloudera-quickstart...
Saved

General

Name: cloudera-quickstart-vm-5.8.0-0-virtualbox
Operating System: Red Hat (64-bit)

System

Base Memory: 8192 MB
Processors: 2
Boot Order: Hard Disk, Optical
Acceleration: VT-x/AMD-V, Nested Paging, PAE/NX, KVM Paravirtualization

Preview

Display

Video Memory: 32 MB
Remote Desktop Server: Disabled
Video Capture: Disabled

Storage

Controller: IDE Controller
IDE Primary Master: cloudera-quickstart-vm-5.8.0-0-virtualbox-disk1.vmdk (Normal, 64.00 GB)
IDE Secondary Master: [Optical Drive] Empty

Audio

Disabled

Network

Adapter 1: Intel PRO/1000 MT Desktop (NAT)

USB

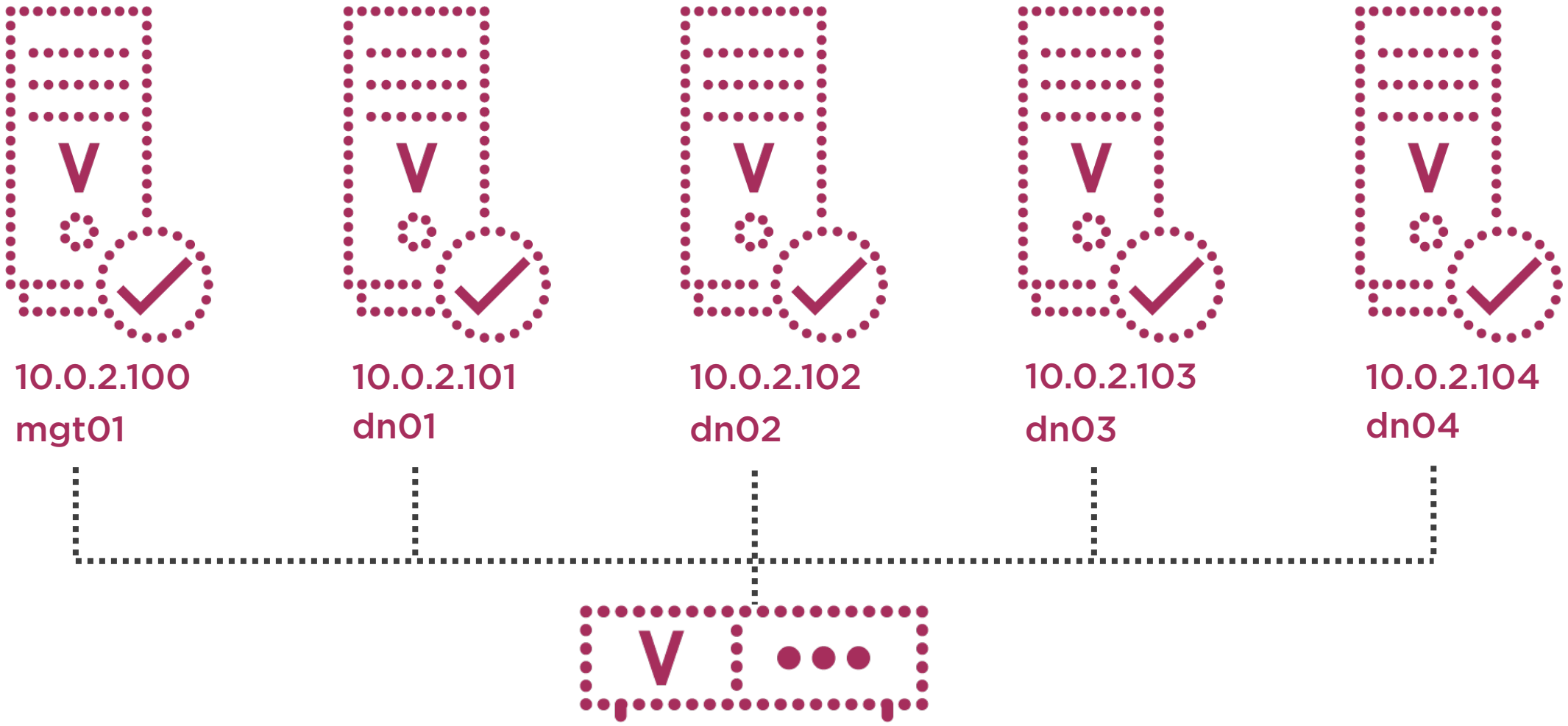
Shared folders

None

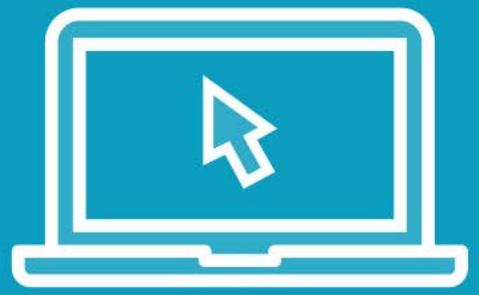
Description

None

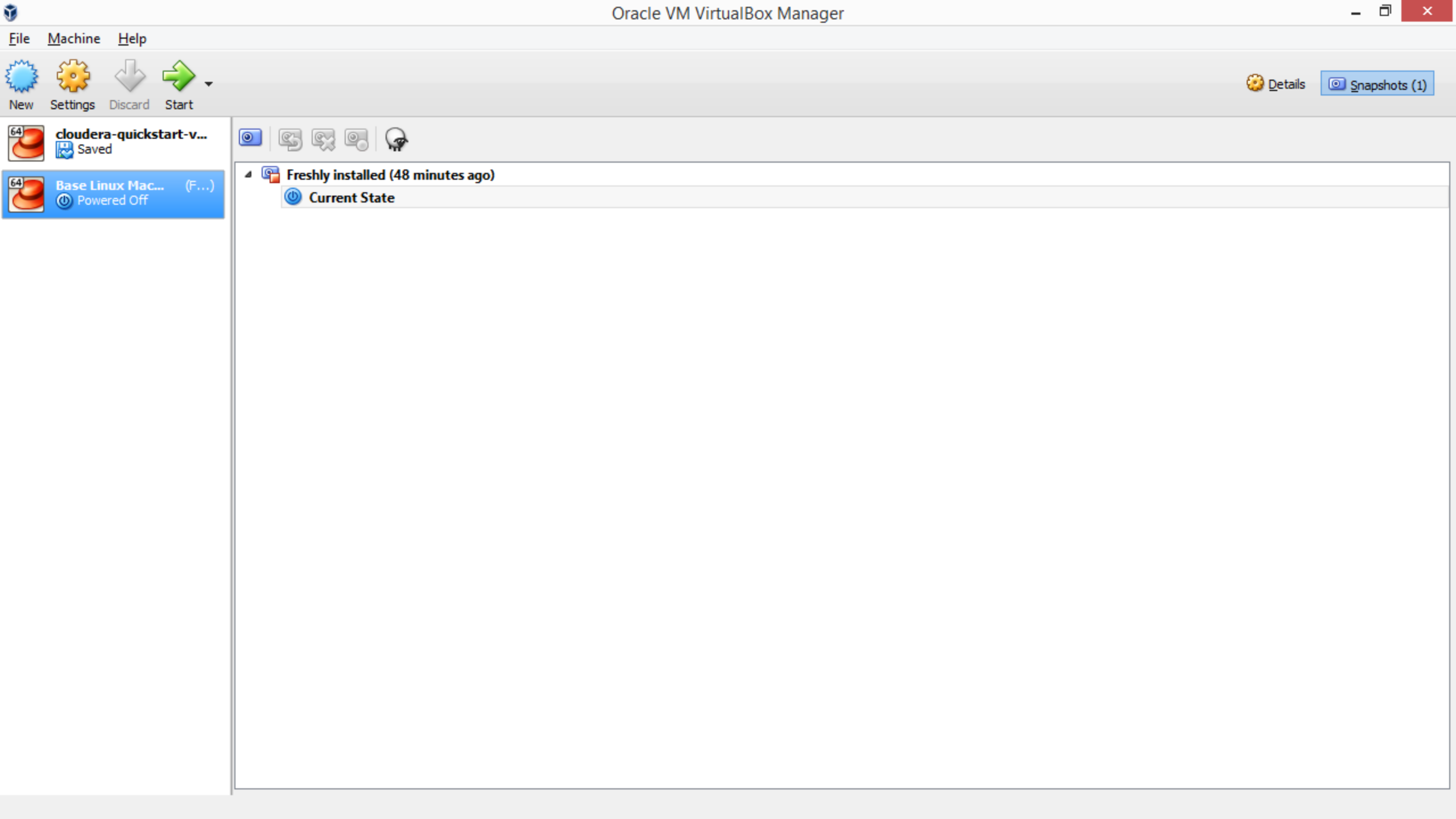
Preparing Your Big Data Linux Cluster



Demo



Exercise 6
Preparing Your Linux Machines for Your Cloudera Big Data Cluster


Oracle VM VirtualBox Manager
File
Machine
Help
New
Settings
Discard
Start
Details
Snapshots (1)
cloudera-quickstart-v...
Saved
Base Linux Mac...
(F...)
Powered Off
Freshly installed (48 minutes ago)
Current State

Takeaway

Linux cluster

CentOS

- Supported by Cloudera

Started from scratch

- Download & install OS
- Configure
- Clone
- Final configurations

Created a base cluster

Ready for Cloudera CDH