

# How to install Hadoop distribution from Bigtop 1.2.0

## Environment

- CentOS 6 machines running on VM/Docker/Bare Metals, etc.
- If using docker, don't run this on the docker host but on a CentOS 6 container

## Deploy

Copy and paste the following text into a file called bigtop-deploy.sh, save the file, chmod it (e.g., `chmod 0744 bigtop-deploy.sh`), and run it (e.g., `./bigtop-deploy.sh`).

Run bigtop-deploy.sh on EVERY node in your cluster:

```

# Welcome
RED='\033[0;31m'; BLUE='\033[0;34m'; DEFAULTC='\033[0m'
printf "\nWelcome to the ${RED}Bigtop${DEFAULTC} installer\n\n"

# Check for CentOS 6
rel=$(cat /etc/redhat-release 2> /dev/null)
if [[ -n "$rel" && ${rel:15:1}==6 ]]; then
    printf "This OS is running ${BLUE}$rel${DEFAULTC} which is compatible with this helper script.\n\n"
else
    printf "This OS is ${RED}not compatible${DEFAULTC} with this Bigtop release. Please install CentOS 6.
x\n\n"
    exit 1
fi

# Get hostname
printf "Enter the hostname of the master node or press Enter to use ${RED}$(hostname -f)${DEFAULTC} : "
read mnode
mnode=${mnode:-$(hostname -f)}
echo

# Install Dependencies
sudo yum -y install git

# Install Puppet
sudo rpm -ivh http://yum.puppetlabs.com/puppetlabs-release-el-6.noarch.rpm
sudo yum -y install puppet
sudo puppet module install puppetlabs-stdlib

# Install Bigtop Puppet
sudo git clone https://github.com/apache/bigtop.git /bigtop-home
sudo sh -c "cd /bigtop-home; git checkout branch-1.2"
sudo cp -r /bigtop-home/bigtop-deploy/puppet/hieradata/ /etc/puppet/
sudo cp /bigtop-home/bigtop-deploy/puppet/hiera.yaml /etc/puppet/

# Configure
sudo su root -c "cat > /etc/puppet/hieradata/site.yaml << EOF
---
bigtop::hadoop_head_node: "$mnode"
hadoop::hadoop_storage_dirs:
- /data/1
- /data/2
hadoop_cluster_node::cluster_components:
- hdfs
- spark
bigtop::jdk_package_name: "java-1.8.0-openjdk-devel.x86_64"
bigtop::bigtop_repo_uri: "http://repos.bigtop.apache.org/releases/1.2.0/centos/6/x86_64"
EOF
"

# Deploy
sudo puppet apply --parser future --modulepath=/bigtop-home/bigtop-deploy/puppet/modules:/etc/puppet/modules
/bigtop-home/bigtop-deploy/puppet/manifests

```

DONE!

## Important settings

- Customize to Hadoop+Spark+HBase Stack

```
hadoop_cluster_node::cluster_components:  
- hdfs  
- yarn  
- spark  
- hbase
```

- Bigtop 1.2.0 CentOS 6 repo:

```
bigtop::bigtop_repo_uri: "http://repos.bigtop.apache.org/releases/1.2.0/centos/6/x86_64"
```

Check out available OS from Bigtop [Installable binary artifacts](#)