

Table of Contents

Crearea environment-ului	1
Prerechizite	1
Crearea workspace-ului vagrant.....	1
Comenzi utile pentru vagrant.....	3
Alte specificatii.....	3
Exercitii.....	3

Crearea environment-ului

Prerechizite

Containerul pentru masinile virtuale folosit este VirtualBox:

<https://www.virtualbox.org/>

Pentru a gestiona masinile virtuale vom folosi Vagrant:

<http://www.vagrantup.com/>

Pentru a descarca workspace-ul de vagrant vom folosi git:

```
sudo apt-get install git
```

Crearea workspace-ului vagrant

Urmati instructiunile din fisierul README.md de pe

<https://github.com/hasnarobert/hadoop-cluster-vagrant/>

Va descarca o imagine cu ubuntu pentru masinile virtuale, va crea 5 masini si va configura hadoop sa ruleze distribuit pe ele.

Deschideti <http://192.168.1.10:50070/> intr-un browser web si trebuie sa va apara urmatoarea pagina:

NameNode 'master:9000'

Started: Wed Oct 23 11:43:26 UTC 2013
Version: 1.2.1, r1503152
Compiled: Mon Jul 22 15:23:09 PDT 2013 by mattf
Upgrades: There are no upgrades in progress.

[Browse the filesystem](#)
[Namenode Logs](#)

Cluster Summary

6 files and directories, 1 blocks = 7 total. Heap Size is 46.11 MB / 966.69 MB (4%)

Configured Capacity : 236.64 GB
DFS Used : 120 KB
Non DFS Used : 19.55 GB
DFS Remaining : 217.09 GB
DFS Used% : 0 %
DFS Remaining% : 91.74 %
Live Nodes : 3
Dead Nodes : 0
Decommissioning Nodes : 0
Number of Under-Replicated Blocks : 0

NameNode Storage:

Storage Directory	Type	State
/tmp/hadoop-root/dfs/name	IMAGE_AND_EDITS	Active



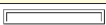
Verificati ca toate DataNode-urile sunt active accesand linkul de "Live Nodes". Trebuie sa va apara urmatoarea pagina:

NameNode 'master:9000'

Started: Wed Oct 23 11:43:26 UTC 2013
Version: 1.2.1, r1503152
Compiled: Mon Jul 22 15:23:09 PDT 2013 by mattf
Upgrades: There are no upgrades in progress.

[Browse the filesystem](#)
[Namenode Logs](#)
[Go back to DFS home](#)

Live Datanodes : 3

Node	Last Contact	Admin State	Configured Capacity (GB)	Used (GB)	Non DFS Used (GB)	Remaining (GB)	Used (%)	Used (%)	Remaining (%)	Blocks
hadoop1	1	In Service	78.88	0	6.52	72.36	0		91.74	1
hadoop2	2	In Service	78.88	0	6.52	72.36	0		91.74	1
hadoop3	1	In Service	78.88	0	6.52	72.36	0		91.74	1

Comenzi utile pentru vagrant

vagrant up <numele-masinii> # creaza/porneste masina virtuala
vagrant destroy <numele-masinii> # distruge masina specificata
vagrant destroy # distruge tot environment-ul
vagrant status # afiseaza starea environment-ului
vagrant ssh <numele-masinii> # Porneste o sesiune ssh catre masina specificata

Alte specificatii

Vagrant creaza 5 masini virtuale: master, backup, hadoop1, hadoop2, hadoop3
Hadoop este instalat in /opt/hadoop-1.2.1

Exercitii

Documentatie hadoop dfs:

http://hadoop.apache.org/docs/r0.18.3/hdfs_shell.html

Copiatii niste fisiere (de preferat destul de mari):

```
bin/hadoop dfs -copyFromLocal /path/to/local/files  
/path/to/hdfs
```

Verificati ca fisierul a fost replicat:

- Spatiul folosit de NameNode din interfata web
- exista o comanda care listeaza blocurile si masinile pe care sunt stocate.
Gasiti-o :D

Opriti o masina cu DataNode vagrant halt hadoopX. Verificati ca fisierele sunt disponibile si contin toata informatia.

Restartati NameNode-ul si verificati starea clusterului.

Modificati factorul de replicare pentru un fisier si verificati modificarile.

Creati un fisier care nu pote fi accesat de utilizatorul curent.

Utilizati cat mai multe dintre comenzile de hdfs. Familiarizati-va cu ele.



■ Cel mai greu exercitiu (ma astept sa nu reuseasca nimeni sa il termine):

adaugati inca un o masina cu un DataNode, atasati-l clusterului **fara sa il opriti**.
Rebalansati blocurile de fisiere astfel incat nodul nou creat sa contina o parte din ele.