**Apache Hadoop** is a collection of open-source software utilitiesthat enables us to deal with Big Data using a network of many computers to solve problems involving massive amounts of data and computation. It’s written in Java

**2018** Apache Hadoop **3.1** Available

**2014** Apache Hadoop **2.5** Available

**The three core components of Hadoop are**

**HDFS:** The storage layer of Hadoop. **Map Reduce: T**he data processing layer of Hadoop based on split-apply-combine strategy. **YARN: T**he resource management layer of Hadoop.

**Features of HDFS:** Distributed, Scalable, Cost-Effective, Fault Tolerant and High Throughput.

**Name Node:** Stores and Manages names of directory and names of files Namespaces. It keeps the directory tree of all files in the file system, and tracks where across the cluster the file data is kept. It does not store the data of these files itself.

**Data Node:** A DN stores actual data in the [Hadoop File System], and meta data for the data stored in block. Metadata is mainly checksum. It periodically sends heartbeat and block report to NN.

**Heart Beat:** Signal sent by the DN to NN on a regular interval to indicate its presence. Default heartbeat interval is 3 seconds defined in hdfs-site.xml file with property. *dfs.heartbeat.interval*

Information contained in Heartbeat are: **Registration:** DN registration information **Capacity:** Total storage capacity available at DN**. dfsUsed:** Storage used by HDFS **remaining:** Remainingstorage available for HDFSblock **PoolUsed:** Storage used by the block pool **xmitsInProgress:** Number of transfers from this DN to other **xceiverCount:** Number of active transceiver. **threadsxmitsInProgress:** Number of transfers from this DN to others. **cacheCapacity:** Total cache capacity available at DN**. cacheUsed:** Amount of cache used.

This information is used to: 1. Mark a DN dead or alive. 2. Registration of new DN. 3 To update Metrics of DN. To issue command to DN. NN can issue following commands to DN. *BlockRecoveryCommand* to recover specified blocks, *BlockCommand* for transferring blocks to another DN, for invalidating certain blocks, *Cache/Uncache* commands for caching / uncaching the blocks

**Block Report:** Block reports interval is controlled by a hdfs-site.xml configuration named as *dfs.blockreport.intervalMsec*. By default, this is set to 21600000 milliseconds i.e. 6 hours. Information contained in Block Report are:registration and block Information, which contains: block ID, block length, block generation timestamp, state of the block replica (For e.g. replica is finalized or waiting to be recovered etc.)

**Block Size:** The default block size in Hadoop 1 is: 64 MB. The default block size in Hadoop 2 is: 128 MB. We can change this size for entire cluster by changing *dfs.block.size* property in hdfs-site.xml. Changing the block size will not affect the block size of any files already in HDFS. It will be applicable for files placed after this setting takes effect.

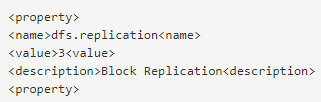


**Replication Factor:** For each block stored in HDFS, there will be n – 1 duplicated blocks distributed across the cluster. Changing the *dfs.replication* property in *hdfs-site.xml* will change default replication for all files placed in HDFS. We can also change the replication factor on a per-file basis using the Hadoop FS shell.

*[myT@localhost ~]$ hadoop fs –setrep –w 3 /my/file*

Alternatively, we can change the replication factor of all the files under a directory.

*[training@localhost ~]$ hadoop fs –setrep –w 3 -R /my/dir*



**Fsimage file:** it contains the complete state of the file system at a point in time. Every file system modification is assigned a unique, monotonically increasing transaction ID. An *fsimage* file represents the file system state after all modifications up to a specific transaction ID. The *fsimage* is read from the disk when NN starts and maintained in memory.

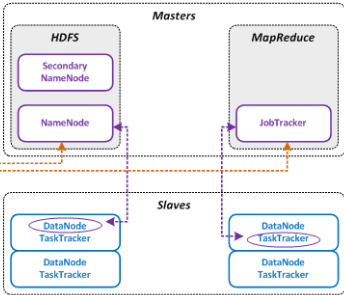
**edits(editlog):** An edits file is a log that lists each file system change (file creation, deletion or modification) that was made after the most recent fsimage. If you like to see your edit logs and fsimage file location, just open the hdfs-site.xml file. *dfs.namenode.name.*[*dir*](https://hadoop.apache.org/docs/r2.4.1/hadoop-project-dist/hadoop-hdfs/hdfs-default.xml) determines where on the local filesystem the DFS NN should store the name table(fsimage).

**Rack Awareness:** In a large cluster of Hadoop, in order to improve the network traffic while reading/writing HDFS file, NN chooses the DN which is closer to the same rack or nearby rack to Read/Write request. NN achieves rack information by maintaining the rack ids of each DN. This concept that chooses closer DN based on the rack information is called Rack Awareness in Hadoop.

**NN and DN Architecture:**

Master Slave architecture one NN and Multiple DN

The NN determines mapping of blocks on DN but after that client directly interact with DN for reading and writing. When client is writing data to HDFS file the data first goes to a local buffer. This approach is adopted to provide streaming read write capability to HDFS.



**Erasure coding:** Let’s assume we have two DNs A and B. To keep data safe in case of failure, we can have replication of A and B, respectively on C and D. That means we are using 4 DNs. But

Erasure coding allows us to achieve the same using just “three” DN. In the 3rd node, we store A+B (actually => A xor B : bit-wise XOR of A and B). The resulting data of (A+B) would have the same size as A or B. Now if node A crashes, we can recover it using the existing data in the 3rd and 2nd node as : (A+B) - B. Reed-Solomon Algorithm– XOR operation limitation is addressed by Reed-Solomon another form of EC

**HDFS Federation** enhances an existing HDFS architecture. In prior HDFS architecture for entire cluster allows only single namespace. In that configuration, Single NN manages namespace. If NN fails, the cluster as a whole would be out of services. The cluster will be unavailable until the NN restarts or brought on a separate machine.

Hadoop Federation overcomes this limitation by adding support for many NN/Namespaces to HDFS.

**Hadoop Schedulers:**

https://data-flair.training/blogs/hadoop-schedulers/

**HDFS Commands:**

* hdfs dfs version
* hdfs dfs -mkdir /user/data/dir1
* hdfs dfs -ls /user/data/dir1
* hdfs dfs put <localSrc> <dest>
* hdfs fs copyFromLocal <localSrc> <dest>
* hdfs fs get [-crc] <src> <localDest>

hdfs dfs -getmerge /user/data/dir2/sample /home/data/Desktop

hadoop fs -getfacl /user/data/dir1/sample

hadoop fs -getfacl -R /user/data/dir1

Above command shows the Access Control Lists (ACLs) of files and directories. If a directory contains a default ACL, then getfacl also displays the default ACL.

* hdfs fs copyToLocal <src> <localDest>
* hdfs dfs -cat /user/data/dir1/sample
* hadoop fs -mv /user/data/dir1/purchases.txt /user/data/dir2
* hadoop fs -cp /user/data/dir2/purchases.txt /user/data/dir1
* hdfs fs moveFromLocal <localSrc> <dest>
* hdfs fs moveToLocal <src> <localDest>
* hdfs dfs -expunge “empty the trash”
* hdfs dfs -chown [-R] [OWNER][:[GROUP]] URI
* hdfs dfs -chgrp [-R] <NewGroupName> <file or directory name>
* hdfs fs setrep [-R] [-w] rep <path>

hdfs dfs -setrep -w 3 /user/data/dir1

The -w flag requests that the command waits for replication process to get completed.  The -R flag is accepted for backward compatibility. It does not make any changes.

* hadoop fs -checksum /user/data/dir1/sample
* hdfs dfs -count /user/data
* hadoop fs -find /user/data/dir1/ -name sam -print
* hadoop fs -help
* hdfs dfs -setfattr -n user.myAttr -v myValue /user/data/dir2/purchases.txt
* hdfs fs du <path>
* hdfs dfs -df [-h] URI [URI ...] hdfs dfs -df -h
* hdfs dfs -touchz /user/dataflair/dir2
* hdfs dfs -test -[ezd] URI
* text hdfs dfs -text <source>
* hadoop fs -appendToFile /file1 /data/dir1
* hdfs dfs -stat path

Bullet Points:

* NN contains the meta data of the entire cluster. It contains the details of each folder, file, replication factor, block names etc
* If heartbeat is not received by NN, then that particular DN is declared as dead.
* Hadoop 2 can be deployed on windows too.
* NN and DN can be installed on same machine for learning purpose.
* hdfs-site.xml file. This file is usually found in the conf/ folder of the Hadoop installation directory

**Introduction to HDFS Disk Balancer**

HDFS provides a command line tool called Disk balancer. It distributes data in a uniform way on all disks of a DN. *dfs.disk.balancer.enabled* in hdfs-site.xml

**Journal Node**: For standby NN to be synchronized with active NN, both communicate with each other through Journal Node. This is useful in configuring Hadoop high Availability cluster using how to configure and manage an HA HDFS cluster, using the Quorum Journal Manager (QJM) feature.

**Different configuration files in Hadoop:**

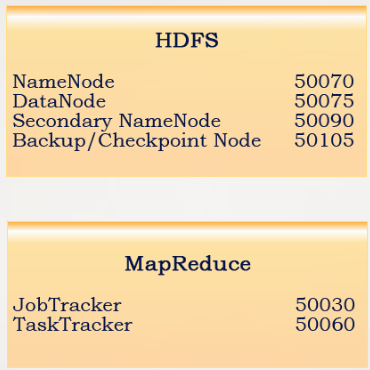
**core-site.xml:** This configuration file contains Hadoop core configuration settings, for example, I/O settings, very common for MapReduce and HDFS. It uses hostname a port.

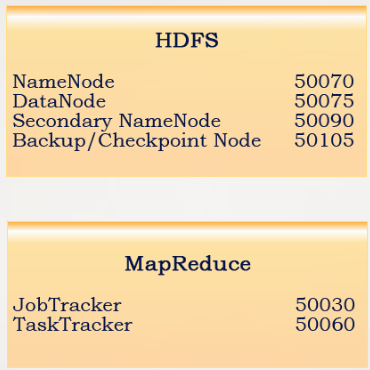
**mapred-site.xml:** This configuration file specifies a framework name for MapReduce by setting mapreduce.framework.name

**hdfs-site.xml:** This configuration file contains HDFS daemons configuration settings. It also specifies default block permission and replication checking on HDFS.

**yarn-site.xml:** This configuration file specifies configuration settings for ResourceManager and NodeManager.

**Port Numbers for NN, Task Tracker, & Job Tracker**





Ambari: http://127.0.0.1:8888/8080

Hortanworks sandbox over Virtualbox in windows.

From terminal fire following ssh client

*ssh 127.0.0.1 –p 2222*

In order to reset password, switch to root

*su root*

*ambari-admin-password-reset*

*ambari-agent restart*

Ways to deal with failure

* Backup Metadata.
* Configure secondary NameNode.
* HDFS Federation and
* HDFS HA Cluster.

HDFS HA Architecture.

Zookeeper keeps track of active NameNode in this case.

Talking to HDFS: Ambari, Command line interface, HTTP/HDFS proxies, JAVA Interface, NFS Gateway.

$ start-dfs.sh, $ stop-dfs.sh

* Create Alert notification in Ambari
* Kerberos security feature.