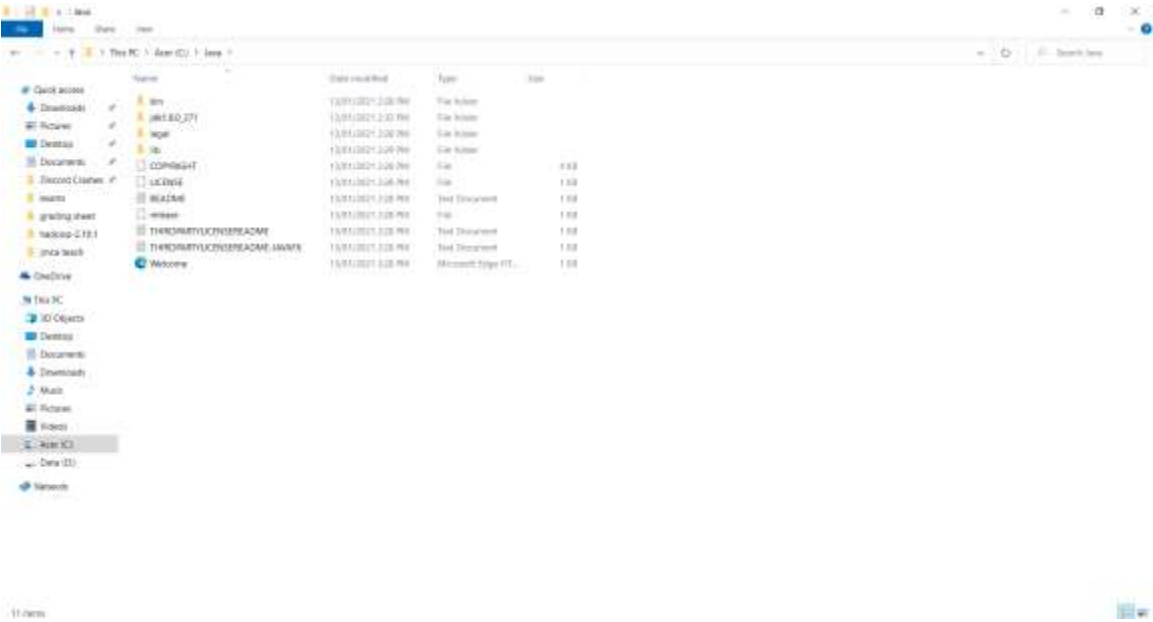
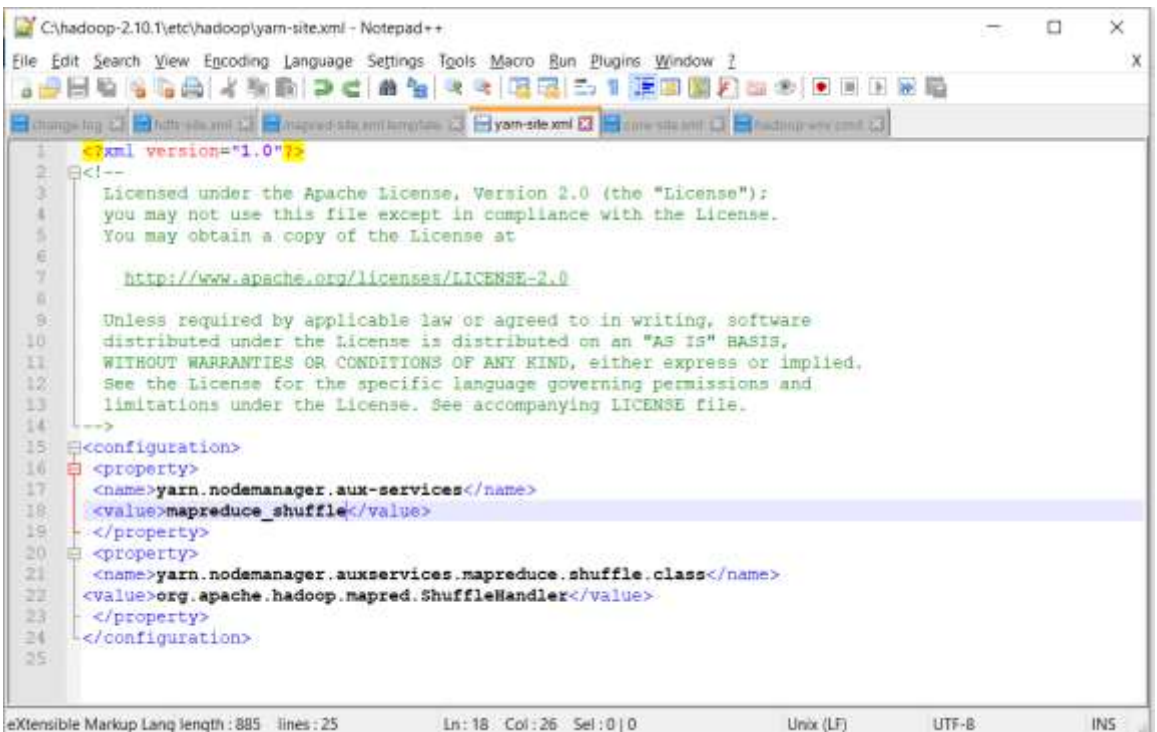
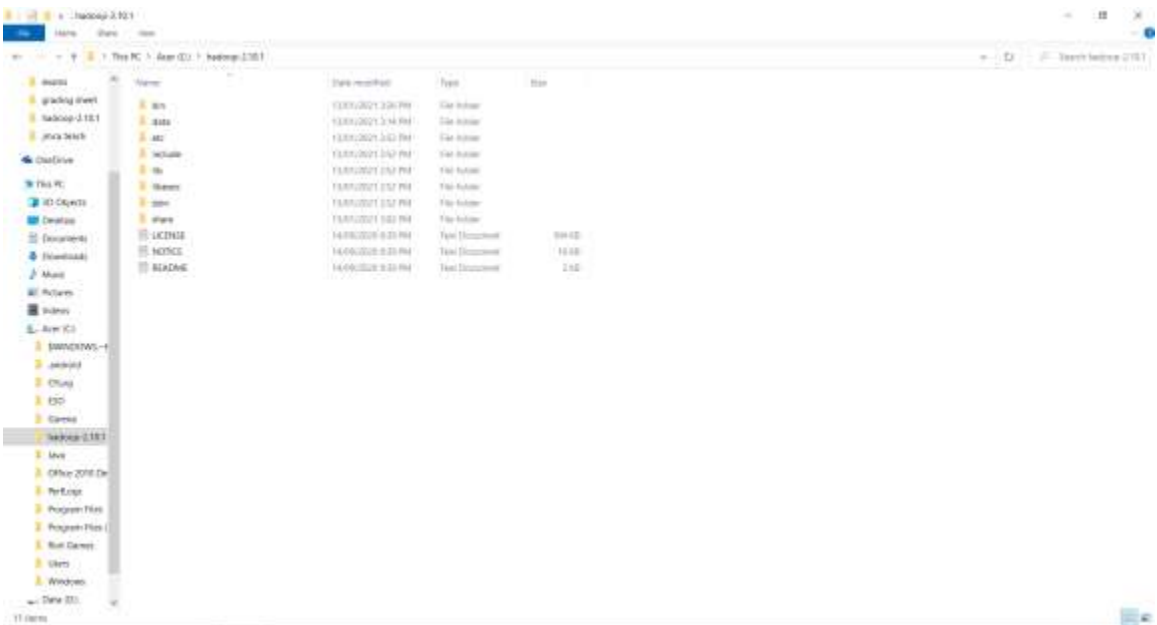


Installing Hadoop and how HDFS architecture (NameNode and DataNode) Works.

First, I downloaded Java



And then I Installed Hadoop (Configuring codes in some of Hadoop XML files: core,mapred,yarn,hdfs,Hadoop env. Codes source from Web) and I created Name Node and Data Node in data Folder and in HDFS XML



CMD CONFIGURE

Checking the Java if it's already installed (Checking the Java Version)

```
Microsoft Windows [Version 10.0.18091.1206]
(c) 2019 Microsoft Corporation. All rights reserved.

C:\Users\A\hadoop>java -version
java version "1.8.0_271"
Java(TM) SE Runtime Environment (build 1.8.0_271-b09)
Java HotSpot(TM) 64-Bit Server VM (build 25.271-b09, mixed mode)
```

Configuring and Creating Namenode (by typing “hdfs name node -format”)

```
Apache Hadoop Distribution - hadoop namenode

STARTUP_MSG:   java = 1.8.0_271
*****/
2020-12-23 16:06:45,316 INFO namenode.NameNode: createNameNode []
2020-12-23 16:06:45,921 INFO impl.MetricsConfig: loaded properties from hadoop-metrics2.properties
2020-12-23 16:06:46,356 INFO impl.MetricsSystemImpl: Scheduled Metric snapshot period at 10 second(s).
2020-12-23 16:06:46,372 INFO impl.MetricsSystemImpl: NameNode metrics system started
2020-12-23 16:06:46,758 INFO namenode.NameNodeUtils: fs.defaultFS is hdfs://localhost:9000
2020-12-23 16:06:46,767 INFO namenode.NameNode: Clients should use localhost:9000 to access this namenode/service.
2020-12-23 16:06:47,124 INFO util.JvmPauseMonitor: Starting JVM pause monitor
2020-12-23 16:06:47,190 INFO hdfs.DFSUtil: Starting Web-server for hdfs at: http://0.0.0.0:9870
2020-12-23 16:06:47,291 INFO util.log: logging initialized @4131ms to org.eclipse.jetty.util.log.Slf4jLog
2020-12-23 16:06:47,636 INFO server.AuthenticationFilter: Unable to initialize FileSignerSecretProvider, falling back to use random secrets.
2020-12-23 16:06:48,040 INFO http.HttpRequestLog: Http request log for http.requests.namenode is not defined
2020-12-23 16:06:48,068 INFO http.HttpServer2: Added global filter 'safety' (class=org.apache.hadoop.http.HttpServer2$QuotingInputFilter)
2020-12-23 16:06:48,072 INFO http.HttpServer2: Added filter static_user_filter (class=org.apache.hadoop.http.lib.StaticUserWebFilter$StaticUserFilter) to context hdfs
2020-12-23 16:06:48,074 INFO http.HttpServer2: Added filter static_user_filter (class=org.apache.hadoop.http.lib.StaticUserWebFilter$StaticUserFilter) to context logs
2020-12-23 16:06:48,075 INFO http.HttpServer2: Added filter static_user_filter (class=org.apache.hadoop.http.lib.StaticUserWebFilter$StaticUserFilter) to context static
2020-12-23 16:06:48,171 INFO http.HttpServer2: Added filter 'org.apache.hadoop.hdfs.web.AuthFilter' (class=org.apache.hadoop.hdfs.web.AuthFilter)
2020-12-23 16:06:48,171 INFO http.HttpServer2: addJerseyResourcePackage: packageName=org.apache.hadoop.hdfs.server.namenode.web.resources;org.apache.hadoop.hdfs.web.resources, pathSpec=/webhdfs/v1/*
2020-12-23 16:06:48,209 INFO http.HttpServer2: Jetty bound to port 9870
2020-12-23 16:06:48,213 INFO server.Server: jetty-9.4.20.v20190813; built: 2019-08-13T21:28:18.144Z; git: 84700530e645e812b336747464d6fbbf370c9a20; jvm 1.8.0_271-b09
2020-12-23 16:06:48,282 INFO server.session: DefaultSessionIdManager workerName=node0
```

After creating Name Node starting all cmd in Hadoop sbin by typing “start-all.cmd”

```
C:\WINDOWS\system32\cmd.exe
Microsoft Windows [Version 10.0.18363.1256]
(c) 2019 Microsoft Corporation. All rights reserved.

C:\Users\W-book>cd \hadoop

C:\hadoop>cd sbin

C:\hadoop\sbin>start-all.cmd
This script is Deprecated. Instead use start-dfs.cmd and start-yarn.cmd
starting yarn daemons

C:\hadoop\sbin>
```

To show and start the Data Node

```
Administrator: Apache Hadoop Distribution
STARTUP_MSG:   java = 1.8.0_271
*****/
2020-12-23 16:06:47,273 INFO checker.ThrottledAsyncChecker: Scheduling a check for [DISK]file:/C:/hadoop-2.8.0/data/datanode
2020-12-23 16:06:47,390 WARN checker.StorageLocationChecker: Exception checking StorageLocation [DISK]file:/C:/hadoop-2.8.0/data/datanode
java.lang.UnsatisfiedLinkError: org.apache.hadoop.io.nativeio.NativeIO$POSIX.stat(Ljava/lang/String;)Lorg/apache/hadoop/io/nativeio/NativeIO$POSIX$Stat;
    at org.apache.hadoop.io.nativeio.NativeIO$POSIX.stat(Native Method)
    at org.apache.hadoop.io.nativeio.NativeIO$POSIX.getStat(NativeIO.java:603)
    at org.apache.hadoop.fs.RawLocalFileSystem$DeprecatedRawLocalFileStatus.loadPermissionInfoByNativeIO(RawLocalFileSystem.java:796)
    at org.apache.hadoop.fs.RawLocalFileSystem$DeprecatedRawLocalFileStatus.loadPermissionInfo(RawLocalFileSystem.java:710)
    at org.apache.hadoop.fs.RawLocalFileSystem$DeprecatedRawLocalFileStatus.getPermission(RawLocalFileSystem.java:678)
    at org.apache.hadoop.util.DiskChecker.mkDirsWithExistsAndPermissionCheck(DiskChecker.java:233)
    at org.apache.hadoop.util.DiskChecker.checkDirInternal(DiskChecker.java:141)
    at org.apache.hadoop.util.DiskChecker.checkDir(DiskChecker.java:116)
    at org.apache.hadoop.hdfs.server.datanode.StorageLocation.check(StorageLocation.java:239)
    at org.apache.hadoop.hdfs.server.datanode.StorageLocation.check(StorageLocation.java:52)
    at org.apache.hadoop.hdfs.server.datanode.checker.ThrottledAsyncChecker$1.call(ThrottledAsyncChecker.java:142)
    at com.google.common.util.concurrent.TrustedListenableFutureTask$TrustedFutureInterruptibleTask.runInterruptibly(TrustedListenableFutureTask.java:125)
    at com.google.common.util.concurrent.InterruptibleTask.run(InterruptibleTask.java:57)
    at com.google.common.util.concurrent.TrustedListenableFutureTask.run(TrustedListenableFutureTask.java:78)
    at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1149)
    at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:624)
    at java.lang.Thread.run(Thread.java:748)
2020-12-23 16:06:47,396 ERROR datanode.DataNode: Exception in secureMain
org.apache.hadoop.util.DiskChecker$DiskErrorException: Too many failed volumes - current valid volumes: 0, volumes configured: 1, volumes failed: 1, volume failures tolerated: 0
```

Resource Manager

```
Administrator: Apache Hadoop Distribution
STARTUP_MSG: java = 1.8.0_271
*****/
2020-12-23 16:06:47,109 INFO conf.Configuration: found resource core-site.xml at file:/C:/hadoop/etc/hadoop/core-site.xml
2020-12-23 16:06:47,267 INFO conf.Configuration: resource-types.xml not found
2020-12-23 16:06:47,287 INFO resource.ResourceUtils: Unable to find 'resource-types.xml'.
2020-12-23 16:06:47,430 INFO conf.Configuration: found resource yarn-site.xml at file:/C:/hadoop/etc/hadoop/yarn-site.xml
2020-12-23 16:06:47,503 INFO event.AsyncDispatcher: Registering class org.apache.hadoop.yarn.server.resourcemanager.RMFatalEventType for class org.apache.hadoop.yarn.server.resourcemanager.ResourceManager$RMFatalEventDispatcher
2020-12-23 16:06:48,204 INFO security.NMTokenSecretManagerInRM: NMTokenKeyRollingInterval: 86400000ms and NMTokenKeyActivationDelay: 900000ms
2020-12-23 16:06:48,215 INFO security.RMContainerTokenSecretManager: ContainerTokenKeyRollingInterval: 86400000ms and ContainerTokenKeyActivationDelay: 900000ms
2020-12-23 16:06:48,230 INFO security.AMRMTokenSecretManager: AMRMTokenKeyRollingInterval: 86400000ms and AMRMTokenKeyActivationDelay: 900000ms
2020-12-23 16:06:48,342 INFO event.AsyncDispatcher: Registering class org.apache.hadoop.yarn.server.resourcemanager.recovery.RMStateStoreEventType for class org.apache.hadoop.yarn.server.resourcemanager.recovery.RMStateStore$ForwardingEventHandler
2020-12-23 16:06:48,357 INFO event.AsyncDispatcher: Registering class org.apache.hadoop.yarn.server.resourcemanager.NodeListManagerEventType for class org.apache.hadoop.yarn.server.resourcemanager.NodeListManager
2020-12-23 16:06:48,358 INFO resource.ResourceManager: Using Scheduler: org.apache.hadoop.yarn.server.resourcemanager.scheduler.capacity.CapacityScheduler
2020-12-23 16:06:48,436 INFO event.AsyncDispatcher: Registering class org.apache.hadoop.yarn.server.resourcemanager.scheduler.event.SchedulerEventType for class org.apache.hadoop.yarn.event.EventDispatcher
2020-12-23 16:06:48,439 INFO event.AsyncDispatcher: Registering class org.apache.hadoop.yarn.server.resourcemanager.rmap.p.RMAppEventType for class org.apache.hadoop.yarn.server.resourcemanager.ResourceManager$ApplicationEventDispatcher
2020-12-23 16:06:48,443 INFO event.AsyncDispatcher: Registering class org.apache.hadoop.yarn.server.resourcemanager.rmap.p.attempt.RMAppAttemptEventType for class org.apache.hadoop.yarn.server.resourcemanager.ResourceManager$ApplicationAttemptEventDispatcher
```

Node Manager

```
Administrator: Apache Hadoop Distribution
STARTUP_MSG: java = 1.8.0_271
*****/
2020-12-23 16:06:47,923 INFO resourceplugin.ResourcePluginManager: No Resource plugins found from configuration!
2020-12-23 16:06:47,923 INFO resourceplugin.ResourcePluginManager: Found Resource plugins from configuration: null
2020-12-23 16:06:48,466 INFO nodemanager.NodeManager: Node Manager health check script is not available or doesn't have execute permission, so not starting the node health script runner.
2020-12-23 16:06:48,593 INFO event.AsyncDispatcher: Registering class org.apache.hadoop.yarn.server.nodemanager.containermanager.container.ContainerEventType for class org.apache.hadoop.yarn.server.nodemanager.containermanager.ContainerManagerImpl$ContainerEventDispatcher
2020-12-23 16:06:48,613 INFO event.AsyncDispatcher: Registering class org.apache.hadoop.yarn.server.nodemanager.containermanager.application.ApplicationEventType for class org.apache.hadoop.yarn.server.nodemanager.containermanager.ContainerManagerImpl$ApplicationEventDispatcher
2020-12-23 16:06:48,646 INFO event.AsyncDispatcher: Registering class org.apache.hadoop.yarn.server.nodemanager.containermanager.localizer.event.LocalizationEventType for class org.apache.hadoop.yarn.server.nodemanager.containermanager.ContainerManagerImpl$LocalizationEventHandlerWrapper
2020-12-23 16:06:48,662 INFO event.AsyncDispatcher: Registering class org.apache.hadoop.yarn.server.nodemanager.containermanager.AuxServicesEventType for class org.apache.hadoop.yarn.server.nodemanager.containermanager.AuxServices
2020-12-23 16:06:48,695 INFO event.AsyncDispatcher: Registering class org.apache.hadoop.yarn.server.nodemanager.containermanager.monitor.ContainersMonitorEventType for class org.apache.hadoop.yarn.server.nodemanager.containermanager.monitor.ContainersMonitorImpl
2020-12-23 16:06:48,737 INFO event.AsyncDispatcher: Registering class org.apache.hadoop.yarn.server.nodemanager.containermanager.launcher.ContainersLauncherEventType for class org.apache.hadoop.yarn.server.nodemanager.containermanager.launcher.ContainersLauncher
2020-12-23 16:06:48,754 INFO event.AsyncDispatcher: Registering class org.apache.hadoop.yarn.server.nodemanager.containermanager.scheduler.ContainerSchedulerEventType for class org.apache.hadoop.yarn.server.nodemanager.containermanager.scheduler.ContainerScheduler
2020-12-23 16:06:48,769 INFO tracker.NMLogAggregationStatusTracker: the rolling interval seconds for the NodeManager Cached Log aggregation status is 600
2020-12-23 16:06:48,880 INFO event.AsyncDispatcher: Registering class org.apache.hadoop.yarn.server.nodemanager.ContainerManagerEventType for class org.apache.hadoop.yarn.server.nodemanager.containermanager.ContainerManagerImpl
```


So, by checking the localhost <http://localhost:9870/> you must not closed all the application nodes so that it won't interrupted.



Finally this will came out and it determines that you downloaded it successfully.

DataNodes usages% (Min/Median/Max/stdDev):	0.00% / 0.00% / 0.00% / 0.00%
Live Nodes	0 (Decommissioned: 0, In Maintenance: 0)
Dead Nodes	0 (Decommissioned: 0, In Maintenance: 0)
Decommissioning Nodes	0
Entering Maintenance Nodes	0
Total Datanode Volume Failures	0 (0 B)
Number of Under-Replicated Blocks	0
Number of Blocks Pending Deletion (including replicas)	0
Block Deletion Start Time	Wed Dec 23 16:06:49 +0800 2020
Last Checkpoint Time	Wed Dec 23 16:06:06 +0800 2020

NameNode Journal Status

Current transaction ID: 1	
Journal Manager	State
FileJournalManager(root=C:\hadoop-2.8.0\data\namenode)	EditLogFileOutputStream(C:\hadoop-2.8.0\data\namenode\current\edits_inprogress_0000000000000000001)

NameNode Storage

Storage Directory	Type	State
C:\hadoop-2.8.0\data\namenode	IMAGE_AND_EDITS	Active

