一、集群部署

	192.168.238.160(xmjmaster)	192.168.238.161(xmjslave1)	192.168.238.162(xmjslave2)
HDFS	NameNode,DataNode	DataNode	datanode,secondarynamenode
YARN	NodeManager	ResourceManager,NodeManager	NodeManager

二、安装Hadoop

```
1.下载
https://archive.apache.org/dist/hadoop/common/hadoop-2.7.2/

2.解压缩
tar -zxvf hadoop-2.7.2.tar.gz -C /usr/local/

3.添加到环境变量
export HADOOP_HOME=/usr/local/hadoop-2.7.2
export PATH=$PATH:$HADOOP_HOME/bin


source /etc/profile
```

三、配置

1.配置文件hadoop-env.sh

```
cd /usr/local/hadoop-2.7.2/etc/hadoop
vim hadoop-env.sh
//文件末尾
export JAVA_HOME=/usr/java/jdk1.8.0_251-amd64
```

2.配置文件：core-site.xml（hdfs的核心配置文件）

```
    <!-- 指定HDFS中NameNode的地址 -->
    <property>
        <name>fs.defaultFS</name>
        <value>hdfs://xmjmaster:9000</value>
    </property>
    <!-- 指定hadoop运行时产生文件的存储目录 -->
    <property>
        <name>hadoop.tmp.dir</name>
        <value>/usr/local/hadoop-2.7.2/data/tmp</value>
    </property>
```

3.hdfs配置文件 hdfs-site.xml

```
    <property>
        <name>dfs.replication</name>
        <value>3</value>
    </property>
    <!--secondarynamenode的地址  辅助namenode工作-->
    <property>
```

```xml
        <name>dfs.namenode.secondary.http-address</name>
        <value>xmjslave2:50090</value>
    </property>
    <property>
        <name>dfs.name.dir</name>
        <value>/home/hadoop/data/hadoop/name/</value>
    </property>
    <property>
        <name>dfs.data.dir</name>
        <value>/home/hadoop/data/hadoop/data/</value>
    </property>
```

4.yarn配置文件yarn-site.xml

```xml
    <!--NodeManager上运行的附属服务。需配置成mapreduce_shuffle，才可运行MapReduce程序>
 -->
    <property>
        <name>yarn.nodemanager.aux-services</name>
        <value>mapreduce_shuffle</value>
    </property>
    <!-- 指定YARN的ResourceManager的地址 -->
    <property>
        <name>yarn.resourcemanager.hostname</name>
        <value>xmjslave1</value>
    </property>
    <property>
        <name>yarn.nodemanager.resource.memory-mb</name>
        <value>256</value>
    </property>
    <property>
        <name>yarn.scheduler.minimum-allocation-mb</name>
        <value>256</value>
    </property>
    <property>
        <name>yarn.scheduler.maximum-allocation-mb</name>
        <value>256</value>
    </property>
    <property>
        <name>yarn.nodemanager.resource.cpu-vcores</name>
        <value>1</value>
    </property>
    <property>
        <name>yarn.scheduler.minimum-allocation-vcores</name>
        <value>1</value>
    </property>
    <property>
        <name>yarn.scheduler.maximum-allocation-vcores</name>
        <value>1</value>
    </property>
```

5.mapreduce配置文件mapred-site.xml

```
cp mapred-site.xml.template mapred-site.xml
vim mapred-site.xml
<!-- 指定mr运行在yarn上 -->
<property>
    <name>mapreduce.framework.name</name>
    <value>yarn</value>
</property>
```

6.配置集群中从节点信息

```
vim slaves
xmjmaster
xmjslave1
xmjslave2
```
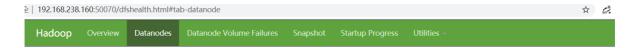
7.将192.168.238.160中hadoop目录下的软件拷贝到其他机器

```
scp -r /usr/local/hadoop-2.7.2 192.168.238.161:/usr/local/
scp -r /usr/local/hadoop-2.7.2 192.168.238.162:/usr/local/
```

四、集群启动测试

1.集群启动

```
1）如果集群是第一次启动，需要格式化NameNode
    hadoop namenode -format
2）整体启动/停止hdfs（在namenode、node节点启动）
    ./sbin/start-dfs.sh
    ./sbin/stop-dfs.sh
3）整体启动/停止yarn （在resourcemanager节点启动）
    ./sbin/start-yarn.sh
    ./sbin/stop-yarn.sh
4）web页面查看
    http://192.168.238.160:50070/
    http://192.168.238.161:8088/
    http://192.168.238.162:50090/status.html
```

| 192.168.238.160:50070/dfshealth.html#tab-datanode | ☆ |

| Hadoop | Overview | Datanodes | Datanode Volume Failures | Snapshot | Startup Progress | Utilities ▾ |

Datanode Information

In operation

Node	Last contact	Admin State	Capacity	Used	Non DFS Used	Remaining	Blocks	Block pool used	Failed Volumes	Version
xmjslave2:50010 (192.168.238.162:50010)	0	In Service	16.99 GB	4 KB	5.22 GB	11.77 GB	0	4 KB (0%)	0	2.7.2
xmjslave1:50010 (192.168.238.161:50010)	0	In Service	16.99 GB	4 KB	6.05 GB	10.94 GB	0	4 KB (0%)	0	2.7.2
xmjmaster:50010 (192.168.238.160:50010)	2	In Service	16.99 GB	4 KB	5.64 GB	11.34 GB	0	4 KB (0%)	0	2.7.2

Logged in as: dr.wh

Nodes of the cluster

Cluster
- About
- Nodes
- Node Labels
- Applications
  - NEW
  - NEW_SAVING
  - SUBMITTED
  - ACCEPTED
  - RUNNING
  - FINISHED
  - FAILED
  - KILLED
- Scheduler

Tools

Cluster Metrics

Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running	Memory Used	Memory Total	Memory Reserved	VCores Used	VCores Total	VCores Reserved	Active Nodes	Decommissioned Nodes	Lost Nodes	Unhealthy Nodes	Rebooted Nodes
0	0	0	0	0	0 B	768 MB	0 B	0	3	0	3	0	0	0	0

Scheduler Metrics

Scheduler Type	Scheduling Resource Type	Minimum Allocation	Maximum Allocation
Capacity Scheduler	[MEMORY]	<memory:256, vCores:1>	<memory:256, vCores:1>

Show 20 entries   Search:

Node Labels	Rack	Node State	Node Address	Node HTTP Address	Last health-update	Health-report	Containers	Mem Used	Mem Avail	VCores Used	VCores Avail	Version
	/default-rack	RUNNING	xmjslave1:35106	xmjslave1:8042	Tue Jul 21 18:37:09 +0800 2020		0	0 B	256 MB	0	1	2.7.2
	/default-rack	RUNNING	xmjslave2:34272	xmjslave2:8042	Tue Jul 21 18:37:06 +0800 2020		0	0 B	256 MB	0	1	2.7.2
	/default-rack	RUNNING	xmjmaster:37095	xmjmaster:8042	Tue Jul 21 18:37:07 +0800 2020		0	0 B	256 MB	0	1	2.7.2

Showing 1 to 3 of 3 entries   First Previous 1 Next Last

```
hadoop fs -mkdir -p /usr/xmj/input
hadoop fs -put README.txt /usr/xmj/input
hadoop fs -ls /usr/xmj/input
```