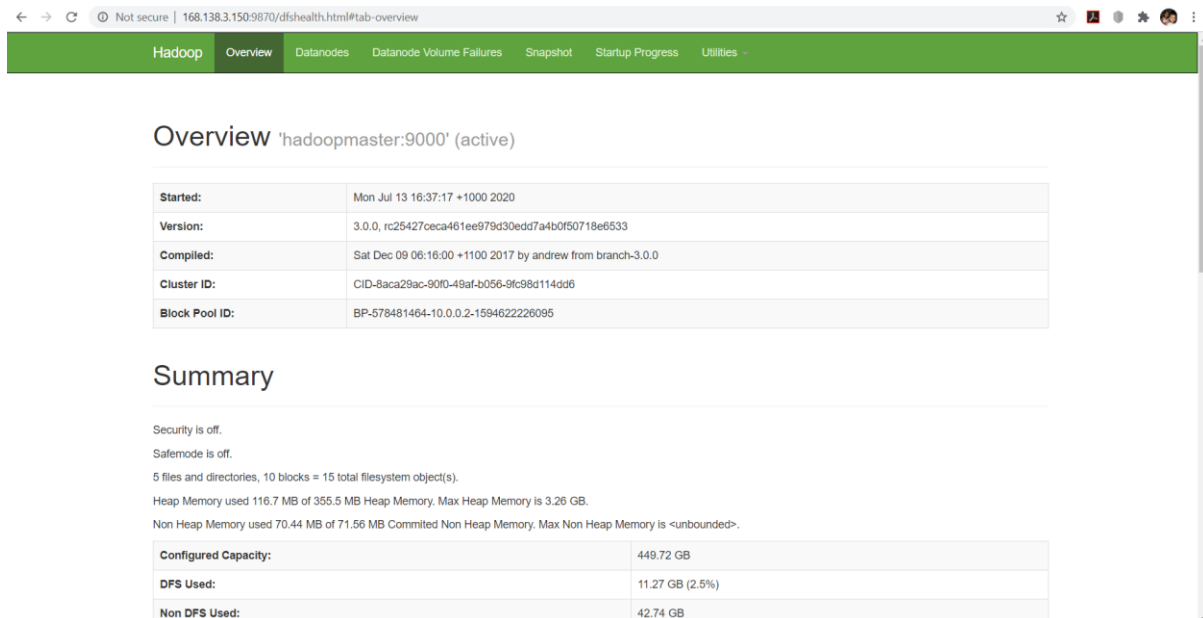


This document lists the common Web Interfaces that the Hadoop services use. These can be used for checking the status of the file system, and troubleshoot issues.

## DFS health on the Name node

<http://masternode-ip:9870>



The screenshot shows the 'Overview' page for the Hadoop NameNode. The browser address bar indicates the URL is `168.138.3.150:9870/dfshealth.html#tab-overview`. The page has a green header with navigation tabs: Hadoop, Overview, Datanodes, Datanode Volume Failures, Snapshot, Startup Progress, and Utilities. The main content area is titled 'Overview 'hadoopmaster:9000' (active)'. It contains a table with key information:

Started:	Mon Jul 13 16:37:17 +1000 2020
Version:	3.0.0, rc25427ceca461ee979d30edd7a4b0f50718e6533
Compiled:	Sat Dec 09 06:16:00 +1100 2017 by andrew from branch-3.0.0
Cluster ID:	CID-8aca29ac-90f0-49af-b056-9fc98d114dd6
Block Pool ID:	BP-578481464-10.0.0.2-159462226095

Below the table is a 'Summary' section with the following text:

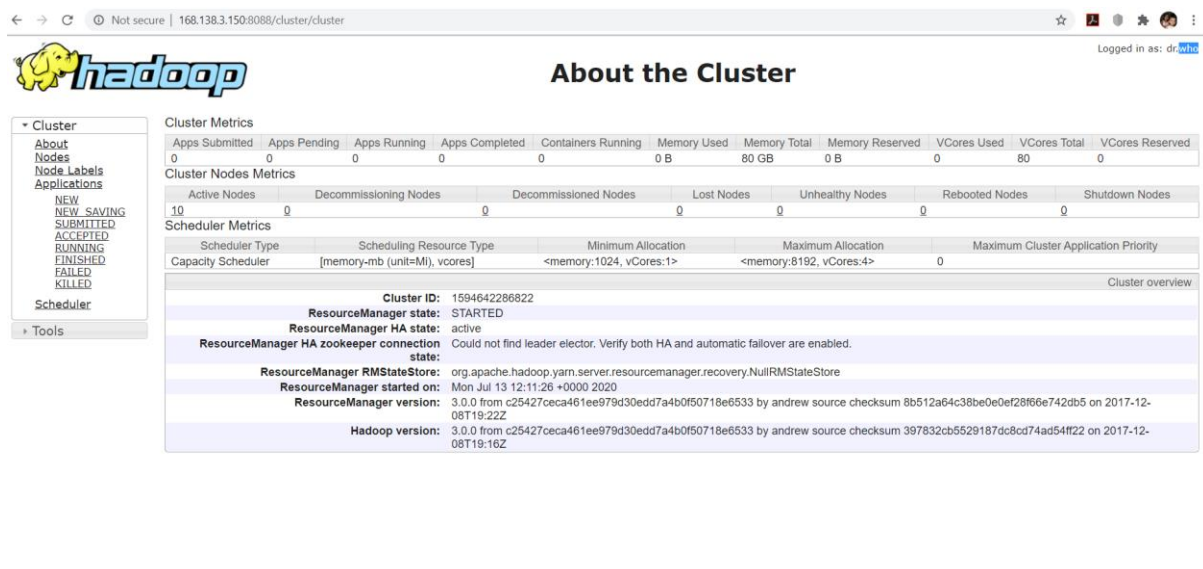
Security is off.  
Safemode is off.  
5 files and directories, 10 blocks = 15 total filesystem object(s).  
Heap Memory used 116.7 MB of 355.5 MB Heap Memory. Max Heap Memory is 3.26 GB.  
Non Heap Memory used 70.44 MB of 71.56 MB Committed Non Heap Memory. Max Non Heap Memory is <unbounded>.

At the bottom, there is a table showing memory usage:

Configured Capacity:	449.72 GB
DFS Used:	11.27 GB (2.5%)
Non DFS Used:	42.74 GB

## Yarn resource manager

<http://masternode-ip:8088>



The screenshot shows the 'About the Cluster' page of the Hadoop Yarn Resource Manager. The browser address bar indicates the URL is `168.138.3.150:8088/cluster/cluster`. The page features the Hadoop logo and a 'Logged in as: dr.who' indicator. A left sidebar contains navigation links: Cluster, About, Nodes, Node Labels, Applications, NEW, NEW SAVING, SUBMITTED, ACCEPTED, RUNNING, FINISHED, FAILED, KILLED, Scheduler, and Tools. The main content area is titled 'About the Cluster' and contains several tables:

**Cluster Metrics**

Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running	Memory Used	Memory Total	Memory Reserved	VCoers Used	VCoers Total	VCoers Reserved
0	0	0	0	0	0 B	80 GB	0 B	0	80	0

**Cluster Nodes Metrics**

Active Nodes	Decommissioning Nodes	Decommissioned Nodes	Lost Nodes	Unhealthy Nodes	Rebooted Nodes	Shutdown Nodes
10	0	0	0	0	0	0

**Scheduler Metrics**

Scheduler Type	Scheduling Resource Type	Minimum Allocation	Maximum Allocation	Maximum Cluster Application Priority
Capacity Scheduler	[memory-mb (unit=Mi), vcores]	<memory:1024, vCores:1>	<memory:8192, vCores:4>	0

**Cluster overview**

Cluster ID:	1594642286822
ResourceManager state:	STARTED
ResourceManager HA state:	active
ResourceManager HA zookeeper connection state:	Could not find leader elector. Verify both HA and automatic failover are enabled.
ResourceManager RMStateStore:	org.apache.hadoop.yarn.server.resourcemanager.recovery.NullRMStateStore
ResourceManager started on:	Mon Jul 13 12:11:26 +0000 2020
ResourceManager version:	3.0.0 from c25427ceca461ee979d30edd7a4b0f50718e6533 by andrew source checksum 8b512a64c38be0e0ef28f66e742db5 on 2017-12-08T19:22Z
Hadoop version:	3.0.0 from c25427ceca461ee979d30edd7a4b0f50718e6533 by andrew source checksum 397832cb5529187dc8d74ad54ff22 on 2017-12-08T19:16Z

## Secondary namenode

<http://masternode-ip:9868>

## Data node health

<http://slavenode-ip:9864>

The screenshot shows the Hadoop DataNode web interface. At the top, there's a green navigation bar with 'Hadoop', 'Overview', and 'Utilities'. Below this, the title 'DataNode on slave6:9866' is displayed. The main content area is divided into three sections: 'Cluster ID' and 'Version' (3.0.0), 'Block Pools' (showing a single pool for hadoopmaster:9000), and 'Volume Information' (showing a single directory for /usr/local/hadoop/hadoopData/hdfs/datanode). The footer indicates 'Hadoop, 2017.'.

Cluster ID:	CID-8aca29ac-90f0-49af-b056-9fc98d114dd6
Version:	3.0.0, rc25427ceca461ee979d30edd7a4b0f50718e6533

Namenode Address	Block Pool ID	Actor State	Last Heartbeat	Last Block Report	Last Block Report Size (Max Size)
hadoopmaster:9000	BP-578481464-10.0.0.2-1594622226095	RUNNING	0s	3 hours	153 B (64 MB)

Directory	Storage Type	Capacity Used	Capacity Left	Capacity Reserved	Reserved Space for Replicas	Blocks
/usr/local/hadoop/hadoopData/hdfs/datanode	DISK	1.13 GB	39.57 GB	0 B	0 B	9

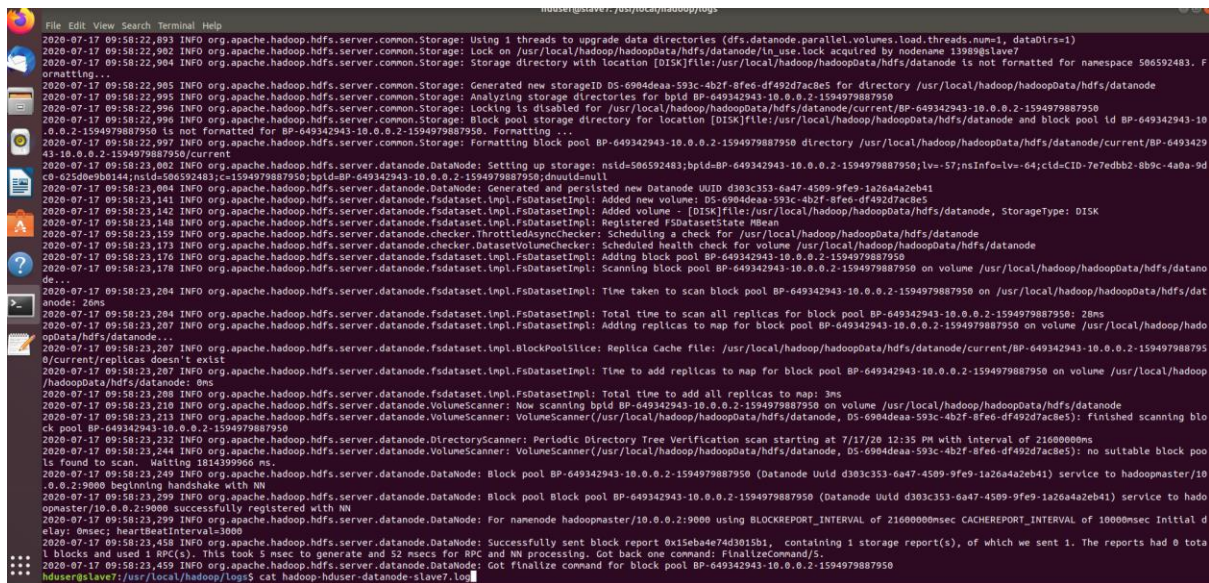
Hadoop, 2017.

These are basically the Web interfaces that we would need while troubleshooting, apart from viewing the logs available at the location:

`/usr/local/hadoop/logs`

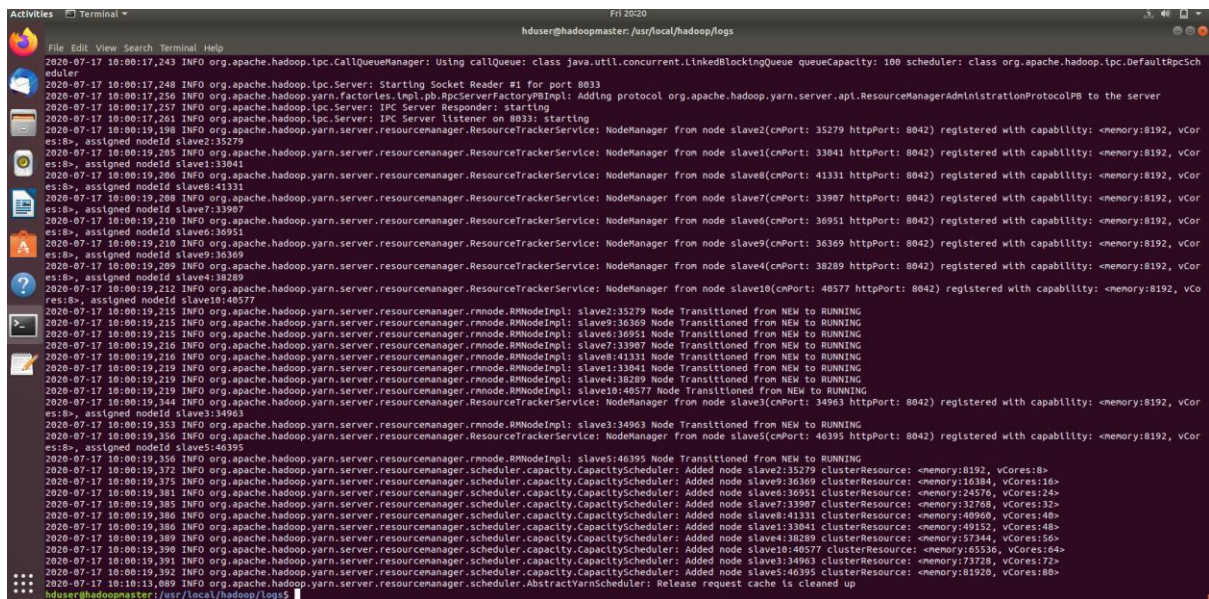
(both master node and the slave nodes).

For example, the below screenshot shows the log of the data node daemon running on slave7:



```
2020-07-17 09:58:22,893 INFO org.apache.hadoop.hdfs.server.common.Storage: Using 1 threads to upgrade data directories (dfs.datanode.parallel.volumes.load.threads.num=1, data0rs=1)
2020-07-17 09:58:22,902 INFO org.apache.hadoop.hdfs.server.common.Storage: Lock on /usr/local/hadoop/hadoopdata/hdfs/datanode/ln_use.lock acquired by nodename 13989@slave7
2020-07-17 09:58:22,904 INFO org.apache.hadoop.hdfs.server.common.Storage: Storage directory with location [DISK]file:/usr/local/hadoop/hadoopdata/hdfs/datanode is not formatted for namespace 506592483. Formatting...
2020-07-17 09:58:22,905 INFO org.apache.hadoop.hdfs.server.common.Storage: Generated new storageID DS-6904dea-593c-4b2f-8fec-df492d7ac8e5 for directory /usr/local/hadoop/hadoopdata/hdfs/datanode
2020-07-17 09:58:22,995 INFO org.apache.hadoop.hdfs.server.common.Storage: Analyzing storage directories for bpid BP-649342943-10.0.0.2-1594979887950
2020-07-17 09:58:22,996 INFO org.apache.hadoop.hdfs.server.common.Storage: Locking is disabled for /usr/local/hadoop/hadoopdata/hdfs/datanode/current/BP-649342943-10.0.0.2-1594979887950
2020-07-17 09:58:22,996 INFO org.apache.hadoop.hdfs.server.common.Storage: Block pool storage directory for location [DISK]file:/usr/local/hadoop/hadoopdata/hdfs/datanode and block pool id BP-649342943-10.0.0.2-1594979887950 is not formatted for BP-649342943-10.0.0.2-1594979887950. Formatting...
2020-07-17 09:58:22,997 INFO org.apache.hadoop.hdfs.server.common.Storage: Formatting block pool BP-649342943-10.0.0.2-1594979887950 directory /usr/local/hadoop/hadoopdata/hdfs/datanode/current/BP-649342943-10.0.0.2-1594979887950
2020-07-17 09:58:23,002 INFO org.apache.hadoop.hdfs.server.datanode.DataNode: Setting up storage: nsid=506592483;bpid=BP-649342943-10.0.0.2-1594979887950;dnuid=null
2020-07-17 09:58:23,004 INFO org.apache.hadoop.hdfs.server.datanode.DataNode: Generated and persisted new Datinode UUID d303c353-6a47-4509-9fe9-1a264a42eb41
2020-07-17 09:58:23,141 INFO org.apache.hadoop.hdfs.server.datanode.FsDatasetImpl: Added new volume: DS-6904dea-593c-4b2f-8fec-df492d7ac8e5
2020-07-17 09:58:23,142 INFO org.apache.hadoop.hdfs.server.datanode.FsDatasetImpl: Added volume - [DISK]file:/usr/local/hadoop/hadoopdata/hdfs/datanode, StorageType: DISK
2020-07-17 09:58:23,148 INFO org.apache.hadoop.hdfs.server.datanode.FsDatasetImpl: Registered FsDatasetState MBean
2020-07-17 09:58:23,159 INFO org.apache.hadoop.hdfs.server.datanode.checker.ThrottledAsyncChecker: Scheduling a check for /usr/local/hadoop/hadoopdata/hdfs/datanode
2020-07-17 09:58:23,173 INFO org.apache.hadoop.hdfs.server.datanode.checker.DatasetVolumeChecker: Scheduled health check for volume /usr/local/hadoop/hadoopdata/hdfs/datanode
2020-07-17 09:58:23,176 INFO org.apache.hadoop.hdfs.server.datanode.FsDatasetImpl: Adding block pool BP-649342943-10.0.0.2-1594979887950 on volume /usr/local/hadoop/hadoopdata/hdfs/datanode
2020-07-17 09:58:23,178 INFO org.apache.hadoop.hdfs.server.datanode.FsDatasetImpl: Scanning block pool BP-649342943-10.0.0.2-1594979887950 on volume /usr/local/hadoop/hadoopdata/hdfs/datanode
2020-07-17 09:58:23,204 INFO org.apache.hadoop.hdfs.server.datanode.FsDatasetImpl: Time taken to scan block pool BP-649342943-10.0.0.2-1594979887950 on /usr/local/hadoop/hadoopdata/hdfs/datanode: 26ms
2020-07-17 09:58:23,204 INFO org.apache.hadoop.hdfs.server.datanode.FsDatasetImpl: Total time to scan all replicas for block pool BP-649342943-10.0.0.2-1594979887950: 26ms
2020-07-17 09:58:23,207 INFO org.apache.hadoop.hdfs.server.datanode.FsDatasetImpl: Adding replicas to map for block pool BP-649342943-10.0.0.2-1594979887950 on volume /usr/local/hadoop/hadoopdata/hdfs/datanode...
2020-07-17 09:58:23,207 INFO org.apache.hadoop.hdfs.server.datanode.FsDatasetImpl: Replica Cache file: /usr/local/hadoop/hadoopdata/hdfs/datanode/current/BP-649342943-10.0.0.2-1594979887950/current/replicas doesn't exist
2020-07-17 09:58:23,207 INFO org.apache.hadoop.hdfs.server.datanode.FsDatasetImpl: Time to add replicas to map for block pool BP-649342943-10.0.0.2-1594979887950 on volume /usr/local/hadoop/hadoopdata/hdfs/datanode: 8ms
2020-07-17 09:58:23,208 INFO org.apache.hadoop.hdfs.server.datanode.FsDatasetImpl: Total time to add all replicas to map: 3ms
2020-07-17 09:58:23,210 INFO org.apache.hadoop.hdfs.server.datanode.VolumesScanner: Now scanning bpid BP-649342943-10.0.0.2-1594979887950 on volume /usr/local/hadoop/hadoopdata/hdfs/datanode
2020-07-17 09:58:23,213 INFO org.apache.hadoop.hdfs.server.datanode.VolumesScanner: VolumesScanner: /usr/local/hadoop/hadoopdata/hdfs/datanode, DS-6904dea-593c-4b2f-8fec-df492d7ac8e5: finished scanning block pool BP-649342943-10.0.0.2-1594979887950
2020-07-17 09:58:23,232 INFO org.apache.hadoop.hdfs.server.datanode.DirectoryScanner: Periodic Directory Tree Verification scan starting at 7/17/20 12:35 PM with interval of 21600000ms
2020-07-17 09:58:23,244 INFO org.apache.hadoop.hdfs.server.datanode.VolumesScanner: VolumesScanner: /usr/local/hadoop/hadoopdata/hdfs/datanode, DS-6904dea-593c-4b2f-8fec-df492d7ac8e5: no suitable block pools found to scan. Halting 1034399956 ms.
2020-07-17 09:58:23,249 INFO org.apache.hadoop.hdfs.server.datanode.DataNode: Block pool BP-649342943-10.0.0.2-1594979887950 (Datinode UUID d303c353-6a47-4509-9fe9-1a264a42eb41) service to hadoopmaster/10.0.0.2:9000 beginning handshake with NN
2020-07-17 09:58:23,299 INFO org.apache.hadoop.hdfs.server.datanode.DataNode: Block pool Block pool BP-649342943-10.0.0.2-1594979887950 (Datinode UUID d303c353-6a47-4509-9fe9-1a264a42eb41) service to hadoopmaster/10.0.0.2:9000 successfully registered with NN
2020-07-17 09:58:23,299 INFO org.apache.hadoop.hdfs.server.datanode.DataNode: For namenode hadoopmaster/10.0.0.2:9000 using BLOCKREPORT_INTERVAL of 21600000msec CACHEREPORT_INTERVAL of 1000000msec Initial delay: 0msec; heartbeatInterval=3000
2020-07-17 09:58:23,458 INFO org.apache.hadoop.hdfs.server.datanode.DataNode: Successfully sent block report 0x15eb4e74d3015b1, containing 1 storage report(s), of which we sent 1. The reports had 0 total blocks and used 1 RPC(s). This took 5 msec to generate and 52 msec for RPC and NN processing. Got back one command: FinalizeCommand/s.
2020-07-17 09:58:23,459 INFO org.apache.hadoop.hdfs.server.datanode.DataNode: Got finalize command for block pool BP-649342943-10.0.0.2-1594979887950
huser@slave7: /usr/local/hadoop/logs$ cat hadoop-hduser-datanode-slave7.log
```

The screenshot below gives the log of Node Manager on the master node:



```
2020-07-17 10:00:17,243 INFO org.apache.hadoop.ipc.CallQueueManager: Using callQueue: class java.util.concurrent.LinkedBlockingQueue queueCapacity: 100 scheduler: class org.apache.hadoop.ipc.DefaultRpcScheduler
2020-07-17 10:00:17,248 INFO org.apache.hadoop.ipc.Server: Starting Socket Reader #1 for port 8033
2020-07-17 10:00:17,256 INFO org.apache.hadoop.yarn.factory.impl.pb.RpcServerFactoryPBImpl: Adding protocol org.apache.hadoop.yarn.server.api.ResourceManagerAdministrationProtocolPB to the server
2020-07-17 10:00:17,261 INFO org.apache.hadoop.ipc.Server: IPC Server Responder: starting
2020-07-17 10:00:17,261 INFO org.apache.hadoop.ipc.Server: IPC Server listener on 8033: starting
2020-07-17 10:00:19,198 INFO org.apache.hadoop.yarn.server.resourcemanager.ResourceTrackerService: NodeManager from node slave2(cnPort: 35279 httpPort: 8042) registered with capability: <memory:8192, vCores:8>, assigned nodeid slave1:33041
2020-07-17 10:00:19,205 INFO org.apache.hadoop.yarn.server.resourcemanager.ResourceTrackerService: NodeManager from node slave1(cnPort: 33041 httpPort: 8042) registered with capability: <memory:8192, vCores:8>, assigned nodeid slave1:33041
2020-07-17 10:00:19,206 INFO org.apache.hadoop.yarn.server.resourcemanager.ResourceTrackerService: NodeManager from node slave8(cnPort: 41331 httpPort: 8042) registered with capability: <memory:8192, vCores:8>, assigned nodeid slave8:41331
2020-07-17 10:00:19,208 INFO org.apache.hadoop.yarn.server.resourcemanager.ResourceTrackerService: NodeManager from node slave7(cnPort: 33907 httpPort: 8042) registered with capability: <memory:8192, vCores:8>, assigned nodeid slave7:33907
2020-07-17 10:00:19,210 INFO org.apache.hadoop.yarn.server.resourcemanager.ResourceTrackerService: NodeManager from node slave9(cnPort: 36369 httpPort: 8042) registered with capability: <memory:8192, vCores:8>, assigned nodeid slave9:36369
2020-07-17 10:00:19,210 INFO org.apache.hadoop.yarn.server.resourcemanager.ResourceTrackerService: NodeManager from node slave4(cnPort: 38289 httpPort: 8042) registered with capability: <memory:8192, vCores:8>, assigned nodeid slave4:38289
2020-07-17 10:00:19,212 INFO org.apache.hadoop.yarn.server.resourcemanager.ResourceTrackerService: NodeManager from node slave10(cnPort: 40577 httpPort: 8042) registered with capability: <memory:8192, vCores:8>, assigned nodeid slave10:40577
2020-07-17 10:00:19,215 INFO org.apache.hadoop.yarn.server.resourcemanager.rmnode.RMNodeImpl: slave2:35279 Node Transitioned from NEW to RUNNING
2020-07-17 10:00:19,215 INFO org.apache.hadoop.yarn.server.resourcemanager.rmnode.RMNodeImpl: slave9:36369 Node Transitioned from NEW to RUNNING
2020-07-17 10:00:19,215 INFO org.apache.hadoop.yarn.server.resourcemanager.rmnode.RMNodeImpl: slave8:41331 Node Transitioned from NEW to RUNNING
2020-07-17 10:00:19,216 INFO org.apache.hadoop.yarn.server.resourcemanager.rmnode.RMNodeImpl: slave7:33907 Node Transitioned from NEW to RUNNING
2020-07-17 10:00:19,216 INFO org.apache.hadoop.yarn.server.resourcemanager.rmnode.RMNodeImpl: slave10:40577 Node Transitioned from NEW to RUNNING
2020-07-17 10:00:19,219 INFO org.apache.hadoop.yarn.server.resourcemanager.rmnode.RMNodeImpl: slave1:33041 Node Transitioned from NEW to RUNNING
2020-07-17 10:00:19,219 INFO org.apache.hadoop.yarn.server.resourcemanager.rmnode.RMNodeImpl: slave4:38289 Node Transitioned from NEW to RUNNING
2020-07-17 10:00:19,219 INFO org.apache.hadoop.yarn.server.resourcemanager.ResourceTrackerService: NodeManager from node slave3(cnPort: 34963 httpPort: 8042) registered with capability: <memory:8192, vCores:8>, assigned nodeid slave3:34963
2020-07-17 10:00:19,353 INFO org.apache.hadoop.yarn.server.resourcemanager.rmnode.RMNodeImpl: slave3:34963 Node Transitioned from NEW to RUNNING
2020-07-17 10:00:19,356 INFO org.apache.hadoop.yarn.server.resourcemanager.ResourceTrackerService: NodeManager from node slave5(cnPort: 46395 httpPort: 8042) registered with capability: <memory:8192, vCores:8>, assigned nodeid slave5:46395
2020-07-17 10:00:19,372 INFO org.apache.hadoop.yarn.server.resourcemanager.rmnode.RMNodeImpl: slave5:46395 Node Transitioned from NEW to RUNNING
2020-07-17 10:00:19,375 INFO org.apache.hadoop.yarn.server.resourcemanager.scheduler.capacity.CapacityScheduler: Added node slave2:35279 clusterResource: <memory:8192, vCores:8>
2020-07-17 10:00:19,381 INFO org.apache.hadoop.yarn.server.resourcemanager.scheduler.capacity.CapacityScheduler: Added node slave9:36369 clusterResource: <memory:10384, vCores:16>
2020-07-17 10:00:19,381 INFO org.apache.hadoop.yarn.server.resourcemanager.scheduler.capacity.CapacityScheduler: Added node slave8:41331 clusterResource: <memory:24376, vCores:24>
2020-07-17 10:00:19,385 INFO org.apache.hadoop.yarn.server.resourcemanager.scheduler.capacity.CapacityScheduler: Added node slave7:33907 clusterResource: <memory:32768, vCores:32>
2020-07-17 10:00:19,386 INFO org.apache.hadoop.yarn.server.resourcemanager.scheduler.capacity.CapacityScheduler: Added node slave10:40577 clusterResource: <memory:49968, vCores:48>
2020-07-17 10:00:19,386 INFO org.apache.hadoop.yarn.server.resourcemanager.scheduler.capacity.CapacityScheduler: Added node slave1:33041 clusterResource: <memory:49152, vCores:48>
2020-07-17 10:00:19,389 INFO org.apache.hadoop.yarn.server.resourcemanager.scheduler.capacity.CapacityScheduler: Added node slave4:38289 clusterResource: <memory:57344, vCores:56>
2020-07-17 10:00:19,390 INFO org.apache.hadoop.yarn.server.resourcemanager.scheduler.capacity.CapacityScheduler: Added node slave10:40577 clusterResource: <memory:65536, vCores:64>
2020-07-17 10:00:19,391 INFO org.apache.hadoop.yarn.server.resourcemanager.scheduler.capacity.CapacityScheduler: Added node slave3:34963 clusterResource: <memory:73728, vCores:72>
2020-07-17 10:00:19,392 INFO org.apache.hadoop.yarn.server.resourcemanager.scheduler.capacity.CapacityScheduler: Added node slave5:46395 clusterResource: <memory:81920, vCores:80>
2020-07-17 10:10:13,089 INFO org.apache.hadoop.yarn.server.resourcemanager.scheduler.AbstractVarnScheduler: Release request cache is cleaned up
huser@hadoopmaster: /usr/local/hadoop/logs$
```

A full list of hadoop3 web services can be found on:

<https://kontext.tech/column/hadoop/265/default-ports-used-by-hadoop-services-hdfs-mapreduce-yarn>