


# Hdfs Experiments

## Introduction

This file contains snapshots of HDFS commands and output in our Hadoop clusters. The experiments were carried in four Case studies focusing mainly on the Uptime and Storage of the namenode metadata.



Logged in as: dr.who

## About the Cluster

Cluster

About Nodes Node Labels ApplicationsNEW NEW\_SAVING SUBMITTED ACCEPTED RUNNING FINISHED FAILED KILLED Scheduler

Tools

Cluster Metrics

Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running	Used Resources	Total Resources	Reserved Resources	Physical Mem Used %	Physical VCores Used %
0	0	0	0	0	<memory:0 B, vCores:0>	<memory:0 B, vCores:0>	<memory:0 B, vCores:0>	0	0

Cluster Nodes Metrics

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

Scheduler Metrics

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

Cluster overview

Cluster ID:

1670269512198

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 Dec 05 13:45:12 -0600 2022

ResourceManager version:

3.2.4 from 7e5d9983b388e372fe640f21f048f2f2ae6e9eba by ubuntu source checksum a8ec9b5946c3975e838a2b608e1278a on 2022-07-12T12:07Z

Hadoop version:

3.2.4 from 7e5d9983b388e372fe640f21f048f2f2ae6e9eba by ubuntu source checksum

## Case 1: (1mb) Files.

## Heap Memory

Hadoop

Overview

Datanodes

Datanode Volume Failures

Snapshot

Startup Progress

Utilities

## Overview 'localhost:9000' (active)

Started:	Tue Dec 06 13:40:28 -0600 2022
Version:	3.2.4, r7e5d9983b388e372fe640f21f048f2f2ae6e9eba
Compiled:	Tue Jul 12 06:58:00 -0500 2022 by ubuntu from branch-3.2.4
Cluster ID:	CID-bd7b5e2d-91c2-49b4-8e03-f02aa099bb43
Block Pool ID:	BP-1493163928-192.168.118.1-1669972225906

## Summary

Security is off.

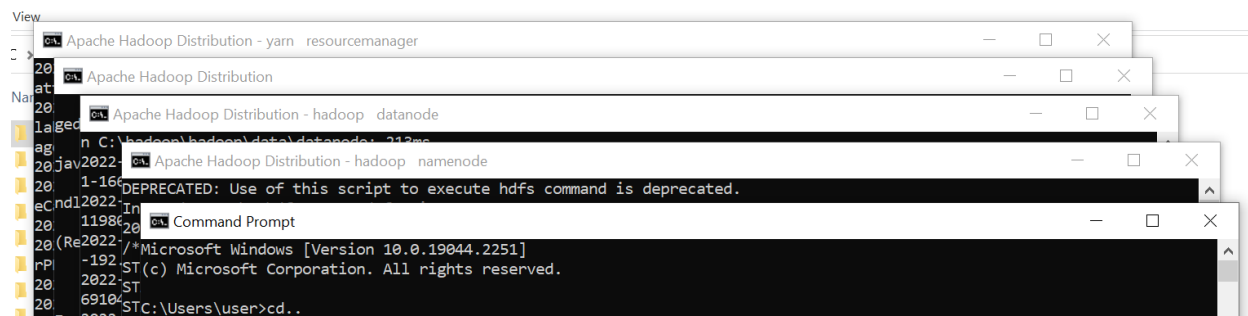
Safemode is off.

1,126 files and directories, 1,122 blocks (1,122 replicated blocks, 0 erasure coded block groups) = 2,248 total filesystem object(s).

Heap Memory used 96.55 MB of 212 MB Heap Memory. Max Heap Memory is 889 MB.

Non Heap Memory used 62.28 MB of 63.81 MB Committed Non Heap Memory. Max Non Heap Memory is <unbounded>.

C:\hadoop\hadoop\data\namenode\current\edits_0000000000000000039-0000000000000006774 1 MB (6736/6736)	100%	
<b>Saving checkpoint</b>	<b>100%</b>	<b>0 sec</b>
erasure coding policies C:\hadoop\hadoop\data\namenode\current\fsimage.ckpt_0000000000000006774 (0/0)	100%	
inodes C:\hadoop\hadoop\data\namenode\current\fsimage.ckpt_0000000000000006774 (0/0)	100%	
delegation tokens C:\hadoop\hadoop\data\namenode\current\fsimage.ckpt_0000000000000006774 (0/0)	100%	
cache pools C:\hadoop\hadoop\data\namenode\current\fsimage.ckpt_0000000000000006774 (0/0)	100%	
<b>Safe mode</b>	<b>100%</b>	<b>33 sec</b>
awaiting reported blocks (1127/1127)	100%	
delegation tokens C:\tmp\hadoop-user\dfs\name\current\fsimage.ckpt_0000000000000008449 (0/0)	100%	
cache pools C:\tmp\hadoop-user\dfs\name\current\fsimage.ckpt_0000000000000008449 (0/0)	100%	
<b>Safe mode</b>	<b>100%</b>	<b>36 sec</b>
awaiting reported blocks (1122/1122)	100%	



```

22:18:08,292 INFO namenode.FSEditLog: Starting log segment at 8449
22:18:08,521 INFO namenode.NameCache: initialized with 0 entries 0 lookups
22:18:08,526 INFO namenode.FSNamesystem: Finished loading FSImage in 2041 msecs
22:18:09,348 INFO namenode.NameNode: RPC server is binding to localhost:9000
22:18:09,348 INFO namenode.NameNode: Enable NameNode state context:false
22:18:09,388 INFO ipc.CallQueueManager: Using callQueue: class java.util.concurrent.LinkedBlockingQueue, queue
1000, scheduler: class org.apache.hadoop.ipc.DefaultRpcScheduler, ipcBackoff: false.

```

```

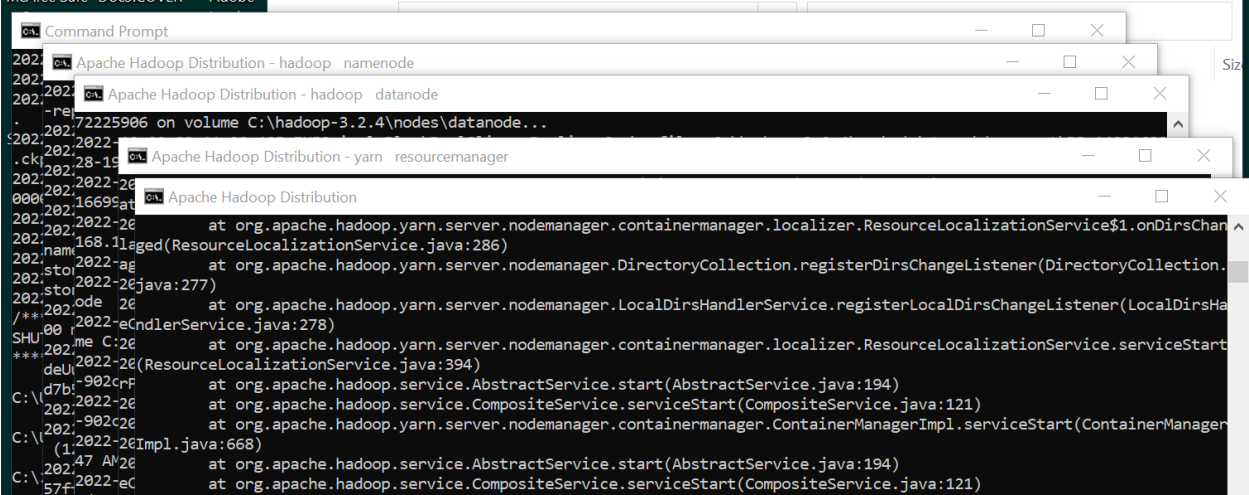
C:\hadoop\hadoop\sbin>hadoop fs -rm -r -f /user/hdfs/1mbfiles
Deleted /user/hdfs/1mbfiles

```

## Killed DataNode

```
2022-12-01 23:09:35,253 INFO ipc.Client: Retrying connect to server: localhost/127.0.0.1:9000. Already tried 2 time(s);
retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECONDS)
2022-12-01 23:09:38,315 INFO ipc.Client: Retrying connect to server: localhost/127.0.0.1:9000. Already tried 3 time(s);
retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECONDS)
2022-12-01 23:09:41,344 INFO ipc.Client: Retrying connect to server: localhost/127.0.0.1:9000. Already tried 4 time(s);
retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECONDS)
2022-12-01 23:09:44,404 INFO ipc.Client: Retrying connect to server: localhost/127.0.0.1:9000. Already tried 5 time(s);
retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECONDS)
2022-12-01 23:09:47,458 INFO ipc.Client: Retrying connect to server: localhost/127.0.0.1:9000. Already tried 6 time(s);
retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECONDS)
2022-12-01 23:09:50,532 INFO ipc.Client: Retrying connect to server: localhost/127.0.0.1:9000. Already tried 7 time(s);
retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECONDS)
2022-12-01 23:09:53,591 INFO ipc.Client: Retrying connect to server: localhost/127.0.0.1:9000. Already tried 8 time(s);
```

Fixed The Datanode



Case 2 (4MB)

Hadoop

Overview

Datanodes

Datanode Volume Failures

Snapshot

Startup Progress

Utilities

Overview 'localhost:9000' (active)

Started:	Tue Dec 06 16:23:37 -0600 2022
Version:	3.2.4, r7e5d9983b388e372fe640f21f048f2f2ae6e9eba
Compiled:	Tue Jul 12 06:58:00 -0500 2022 by ubuntu from branch-3.2.4
Cluster ID:	CID-bd7b5e2d-91c2-49b4-8e03-f02aa099bb43
Block Pool ID:	BP-1493163928-192.168.118.1-1669972225906

Summary

Security is off.

Safemode is off.

261 files and directories, 257 blocks (257 replicated blocks, 0 erasure coded block groups) = 518 total filesystem object(s).

Heap Memory used 65.06 MB of 243.5 MB Heap Memory. Max Heap Memory is 889 MB.

Non Heap Memory used 58.4 MB of 59.73 MB Committed Non Heap Memory. Max Non Heap Memory is <unbounded>.

```
2022-12-02 17:07:57,568 INFO namenode.FSEditLog: Starting log segment at 1580
2022-12-02 17:07:57,725 INFO namenode.NameCache: initialized with 0 entries 0 lookups
2022-12-02 17:07:57,726 INFO namenode.FSNamesystem: Finished loading FSImage in 1399 msecs
2022-12-02 17:07:58,323 INFO namenode.NameNode: RPC server is binding to localhost:9000
2022-12-02 17:07:58,324 INFO namenode.NameNode: Enable NameNode state context:false
2022-12-02 17:07:58,364 INFO ipc.CallQueueManager: Using callQueue: class java.util.concurrent.LinkedBlockingQueue, queueSize: 1000, scheduledThreadPoolName: namenode-kernel-io, DefaultThreadPoolName: DefaultThreadPool
```

Case 3 (154MB)

```
C:\hadoop-3.2.4\sbin>hadoop fs -put C:\Data\179mbfiles /user/hdfs

C:\hadoop-3.2.4\sbin>hadoop fs -ls /user/hdfs
Found 1 items
drwxr-xr-x - user supergroup          0 2022-12-02 20:30 /user/hdfs/179mbfiles
```

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

Overview 'localhost:9000' (active)

Started:	Tue Dec 06 14:50:00 -0600 2022
Version:	3.2.4, r7e5d9983b388e372fe640f21f048f2f2ae6e9eba
Compiled:	Tue Jul 12 06:58:00 -0500 2022 by ubuntu from branch-3.2.4
Cluster ID:	CID-bd7b5e2d-91c2-49b4-8e03-f02aa099bb43
Block Pool ID:	BP-1493163928-192.168.118.1-1669972225906

Summary

Security is off.  
Safemode is off.  
11 files and directories, 13 blocks (13 replicated blocks, 0 erasure coded block groups) = 24 total filesystem object(s).  
Heap Memory used 54.17 MB of 211.5 MB Heap Memory. Max Heap Memory is 889 MB.

```
2022-12-02 21:37:26,148 INFO namenode.FSEditLog: Starting log segment at 1713
2022-12-02 21:37:26,354 INFO namenode.NameCache: initialized with 0 entries 0 lookups
2022-12-02 21:37:26,355 INFO namenode.FSNamesystem: Finished loading FSImage in 941 msecs
2022-12-02 21:37:27,033 INFO namenode.NameNode: RPC server is binding to localhost:9000
2022-12-02 21:37:27,035 INFO namenode.NameNode: Enable NameNode state context:false
```

cache pools (0/0)	100%	
Loading edits	100%	0 sec
\tmp\hadoop-user\dfs\name\current\edits_0000000000000001582-0000000000000001643 1 MB (62/62)	100%	
Saving checkpoint	100%	0 sec
Safe mode	100%	36 sec
awaiting reported blocks (13/13)	100%	

Case 4: 1082MB (One File)

```
C:\hadoop\hadoop\sbin>hadoop fs -ls /user/hdfs/
Found 1 items
drwxr-xr-x - user supergroup          0 2022-12-01 22:37 /user/hdfs/1gigfile

C:\hadoop\hadoop\sbin>
```

```
2022-12-02 20:36:16,169 INFO namenode.FSEditLog: Starting log segment at 1644
2022-12-02 20:36:16,266 INFO namenode.NameCache: initialized with 0 entries 0 lookups
2022-12-02 20:36:16,266 INFO namenode.FSNamesystem: Finished loading FSImage in 478 msec
2022-12-02 20:36:16,538 INFO namenode.NameNode: RPC server is binding to localhost:9000
2022-12-02 20:36:16,538 INFO namenode.NameNode: Enable NameNode state context:false
2022-12-02 20:36:16,551 INFO ipc.CallQueueManager: Using callQueue: class java.util.concurrent.LinkedBlockingQueue, queueCapacity: 1000, scheduler: class org.apache.hadoop.ipc.DefaultRpcScheduler, ipcBackoff: false.
```

# Overview 'localhost:9000' (active)

Started:	Tue Dec 06 15:07:45 -0600 2022
Version:	3.2.4, r7e5d9983b388e372fe640f21f048f2f2ae6e9eba
Compiled:	Tue Jul 12 06:58:00 -0500 2022 by ubuntu from branch-3.2.4
Cluster ID:	CID-bd7b5e2d-91c2-49b4-8e03-f02aa099bb43
Block Pool ID:	BP-1493163928-192.168.118.1-1669972225906

# Summary

Security is off.

Safemode is off.

5 files and directories, 9 blocks (9 replicated blocks, 0 erasure coded block groups) = 14 total filesystem object(s).

Heap Memory used 37.82 MB of 201.5 MB Heap Memory. Max Heap Memory is 889 MB.

```
2022-12-02 16:55:16,995 INFO snapshot.snapshotmanager: Snapshot is disabled
2022-12-02 16:55:16,000 INFO util.GSet: Computing capacity for map cachedBlocks
2022-12-02 16:55:16,000 INFO util.GSet: VM type = 64-bit
2022-12-02 16:55:16,011 INFO util.GSet: 0.25% max memory 889 MB = 2.2 MB
2022-12-02 16:55:16,014 INFO util.GSet: capacity = 2^18 = 262144 entries
2022-12-02 16:55:16,055 INFO metrics.TopMetrics: NNTop conf: dfs.namenode.top.window.num.buckets = 10
2022-12-02 16:55:16,056 INFO metrics.TopMetrics: NNTop conf: dfs.namenode.top.num.users = 10
2022-12-02 16:55:16,059 INFO metrics.TopMetrics: NNTop conf: dfs.namenode.top.windows.minutes = 1,5,25
2022-12-02 16:55:16,092 INFO namenode.FSNamesystem: Retry cache on namenode is enabled
```