# HiBench - How to Check HDFS Configuration and Fix hibench.hdfs.master

## Step 1. Check hdfs-site.xml

In your cluster nodes, open /etc/hadoop/conf/hdfs-site.xml and verify:

<property>  
 <name>dfs.nameservices</name>  
 <value>mycluster</value>  
</property>

<property>  
 <name>dfs.ha.namenodes.mycluster</name>  
 <value>nn1,nn2</value>  
</property>

<property>  
 <name>dfs.namenode.rpc-address.mycluster.nn1</name>  
 <value>hostname1:8020</value>  
</property>

<property>  
 <name>dfs.namenode.rpc-address.mycluster.nn2</name>  
 <value>hostname2:8020</value>  
</property>

If these exist, the correct HDFS URI for HiBench is: hdfs://mycluster  
Never use IPs like hdfs://10.x.x.x or localhost when HA is configured.

## Step 2. Check core-site.xml

Open /etc/hadoop/conf/core-site.xml and verify:

<property>  
 <name>fs.defaultFS</name>  
 <value>wasbs://bucket@storageaccount.blob.core.windows.net</value>  
</property>

If fs.defaultFS is WASBS (Azure Blob), this does NOT affect HiBench if you override hibench.hdfs.master manually.

## Step 3. (Optional) Export Variables Temporarily

To force hdfs://mycluster manually in the current terminal session:

export HADOOP\_CONF\_DIR=/etc/hadoop/conf  
export HADOOP\_CLIENT\_OPTS="-Dfs.defaultFS=hdfs://mycluster"

This ensures any Hadoop command uses HDFS correctly.

## Step 4. Fix HiBench/conf/hadoop.conf

Edit the file HiBench/conf/hadoop.conf:

hibench.hadoop.home /usr/hdp/current/hadoop-client  
hibench.hadoop.executable ${hibench.hadoop.home}/bin/hadoop  
hibench.hadoop.configure.dir /etc/hadoop/conf  
hibench.hdfs.master hdfs://mycluster  
hibench.hadoop.release apache  
hibench.hadoop.examples.jar /usr/hdp/current/hadoop-mapreduce-client/hadoop-mapreduce-examples.jar

Now HiBench will correctly generate input/output paths using HDFS.

## Summary Checklist

| Step | Action | Goal |  
|:---|:---|:---|  
| 1 | Check hdfs-site.xml | Confirm nameservice and HA setup |  
| 2 | Check core-site.xml | Confirm WASBS, ignore it |  
| 3 | Export HADOOP\_CLIENT\_OPTS manually (optional) | Force correct HDFS use in shell |  
| 4 | Fix hibench.hdfs.master in hadoop.conf | Force HiBench to use hdfs://mycluster |

## Optional Mini Script for Automation

#!/bin/bash  
export HADOOP\_CONF\_DIR=/etc/hadoop/conf  
export HADOOP\_CLIENT\_OPTS="-Dfs.defaultFS=hdfs://mycluster"  
echo "Environment exported for HiBench: HDFS Master set to hdfs://mycluster"

## END

After these steps, HiBench will work properly over HDFS without errors like UnknownHostException or Connection Refused.  
No need to modify HDInsight configurations permanently.