TroubleShooting

Some problems encountered in Hadoop and ways to go about solving them. See also NameNodeFailover and ConnectionRefused.

NameNode startup fails

Exception when initializing the filesystem

```
ERROR org.apache.hadoop.dfs.NameNode: java.io.EOFException
    at java.io.DataInputStream.readFully(DataInputStream.java:178)
    at org.apache.hadoop.io.UTF8.readFields(UTF8.java:106)
    at org.apache.hadoop.io.ArrayWritable.readFields(ArrayWritable.java:90)
    at org.apache.hadoop.dfs.FSEditLog.loadFSEdits(FSEditLog.java:433)
    at org.apache.hadoop.dfs.FSImage.loadFSEdits(FSImage.java:759)
    at org.apache.hadoop.dfs.FSImage.loadFSImage(FSImage.java:639)
    at org.apache.hadoop.dfs.FSImage.recoverTransitionRead(FSImage.java:222)
    at org.apache.hadoop.dfs.FSDirectory.loadFSImage(FSDirectory.java:79)
    at org.apache.hadoop.dfs.FSNamesystem.initialize(FSNamesystem.java:254)
    at org.apache.hadoop.dfs.FSNamesystem.<init>(FSNamesystem.java:235)
    at org.apache.hadoop.dfs.NameNode.initialize(NameNode.java:131)
    at org.apache.hadoop.dfs.NameNode.<init>(NameNode.java:176)
    at org.apache.hadoop.dfs.NameNode.<init>(NameNode.java:162)
    at org.apache.hadoop.dfs.NameNode.createNameNode(NameNode.java:846)
    at org.apache.hadoop.dfs.NameNode.main(NameNode.java:855)
```

This is sometimes encountered if there is a corruption of the

```
edits
```

file in the transaction log. Try using a hex editor or equivalent to open up 'edits' and get rid of the last record. In all cases, the last record might not be complete so your NameNode is not starting. Once you update your edits, start the NameNode and run

```
hadoop fsck /
```

to see if you have any corrupt files and fix/get rid of them.

Take a back up of

```
dfs.name.dir
```

before updating and playing around with it.

Client cannot talk to filesystem

Network Error Messages

- BindException
- ConnectionRefused
- EOFException
- NoRouteToHost
- SocketTimeout
- UnknownHost

Error message: Could not get block locations. Aborting...

There are a number of possible of causes for this.

- The NameNode may be overloaded. Check the logs for messages that say "discarding calls..."
- There may not be enough (any) DataNode nodes running for the data to be written. Again, check the logs.
- Every DataNode on which the blocks were stored might be down (or not connected to the NameNode; it is impossible to distinguish the two).

Error message: Could not obtain block

Your logs contain something like

```
INFO hdfs.DFSClient: Could not obtain block blk_-4157273618194597760_1160 from any node:
 java.io.IOException: No live nodes contain current block
```

There are no live DataNode nodes containing a copy of the block of the file you are looking for. Bring up any nodes that are down, or skip that block.

Reduce hangs

This can be a DNS issue. Two problems which have been encountered in practice are:

- Machines with multiple NICs. In this case, set
  ```
  dfs.datanode.dns.interface
  ```
  (in
  ```
  hdfs-site.xml
  ```
) and
  ```
  mapred.datanode.dns.interface
  ```
  (in
  ```
  mapred-site.xml
  ```
) to the name of the network interface used by Hadoop (something like
  ```
  eth0
  ```
  under Linux),
- Badly formatted or incorrect hosts and DNS files (
  ```
  /etc/hosts
  ```
  and {
  ```
  /etc/resolv.conf
  ```
  under Linux) can wreak havoc. Any DNS problem will hobble Hadoop, so ensure that names can be resolved correctly.

Error message saying a file "Could only be replicated to 0 nodes instead of 1"

(or any similar number such as "2 nodes instead of 3")

See CouldOnlyBeReplicatedTo

Client unable to connect to server, "Server not available"

See ServerNotAvailable.

Error message : Too Many Open Files on client or server

See TooManyOpenFiles