1. Default block size

128mb

1. Namenode vs datanode

Namenode holds metadata of all data nodes, data nodes carry out tasks assigned by name node. Data nodes sends heartbeat to namenode every 10 secs. If no signal for 3 mins it will declare it dead

1. How namenode handle datanode failures

Using heartbeats

1. What is metadata and where is it stored

Inside namenode. Metadata is the information regarding all the data

1. How to restart namenode and its deamons
2. How many copies by default/ how to change it

3 copies by default

1. What if block is corrupted
2. How read/write operations work in hdfs
3. How frequent heartbeats are sent

Every 10 seconds

1. Secondary namenode

Its a backup for namenode..it holds all the data that is present in namenode

1. Safemode
2. Mapreduce input split vs hdfs block
3. What happenes if there are too many small files in hdfs
4. HDFS vs NAS
5. How to create users
6. What happens when namenode goes down while writing a file
7. Write once read many
8. Ideal block size for optimization
9. Indexing
10. Rack awareness
11. Commands used in hdfs

Hdfs fs -ls

-cp

-mv

-copyToLocal

-put

-copyFromLocal

Journal node

Job tracker/resource manager

Task tracker/ node manager

Secondary namenode

High availability

Passive namenode

https://data-flair.training/blogs/top-hdfs-interview-questions-and-answers/