

Hadoop Interview Questions

- Q1.) Explain in detail about Kafka Producer in context to Hadoop?
- Q2.) Explain Monad class?
- Q3.) Explain the reliability of Flume-NG data?
- Q4.) What is Interceptor?
- Q5.) What are the different Flume-NG Channel types?
- Q6.) What is Base class in java?
- Q7.) What is Base class in scala?
- Q8.) What is Resilient Distributed Dataset(RDD)?
- Q9.) Give a brief description of Fault tolerance in Hadoop?
- Q10.) What is Immutable data with respect to Hadoop?
- Q11.) Which are the nodes that hadoop can be executed?
- Q12.) How is formatting done in HDFS?
- Q13.) What are the contents found in masterfile of hadoop?
- Q14.) Describe the main hdfs-site.xml properties?
- Q15.) Explain about spill factor with respect to the RAM?
- Q16.) Why do we require a password-less SSH in Fully Distributed environment?
- Q17.) Does this requirement lead to security issues?
- Q18.) What will happen to a NameNode, when ResourceManager is down?
- Q19.) 1 Tell about features of Fully Distributed mode?
- Q20.) Explain about fsck?
- Q21.) how to copy file from local hard disk to hdfs
- Q22.) is it possible to set the reducer to zero???
- Q23.) map-side join / hive join
- Q24.) Managed Table Vs External Table
- Q25.) Difference between bucketing and partitioning

- Q26.) Syntax to create hive table with partitioning
- Q27.) SQOOP split by:
- Q28.) file formats available in SQOOP Import
- Q29.) Default number of mappers in a sqoop command
- Q30.) Maximum number of mappers used a sqoop import command
- Q31.) Flume Architecture
- Q32.) In Unix, command to show all processes
- Q33.) partitions in hive
- Q34.) File formats in hive
- Q35.) Syntax to create bucketed table
- Q36.) Custom Partitioning
- Q37.) Difference between order by and sort by
- Q38.) Purpose of Zoo Keeper
- Q39.) Sqoop Incremental last modified
- Q40.) Difference MR1 vs MR2
- Q41.) select * from table – give what results for normal table and partitioned table
- Q42.) Explode and implode in hive
- Q43.) Interceptors in Flume:
- Q44.) Different types of distributed file systems:
- Q45.) Write a pig script to extract hive table
- Q46.) predefined value in sqoop to extract data from any database current date minus one
- Q47.) UNION, UNIONALL, MINUS and INTERSECT available in hive ?
- Q48.) Difference between Distribute by, cluster by, order by, sort by
- Q49.) Describe the main hdfs-site.xml properties?
- Q50.) What is Hadoop? Name the Main Parts of a Hadoop Application
- Q51.) How many Data Formats are there in Hadoop?
- Q52.)What do you k now about YARN?

- Q53.) Why do nodes are extracted and added regularly in Hadoop cluster?
- Q54.) What do you understand by “Rack Awareness”?
- Q55.) What do you understand about the Speculative Execution?
- Q56.) State any of the main features of Hadoop
- Q57.) Do you know some organizations that are using Hadoop?
- Q58.) How can you distinguish RDBMS RDBMS and Hadoop?
- Q59.) What do you know about active and passive NameNodes?
- Q60.) What are the Parts from Apache HBase?
- Q61.) How is the DataNode failure managed by NameNode?
- Q62.) Define the NameNode recovery process.
- Q63.) What are the various programs available in Hadoop?
- Q64.) Can DataNode and NameNode be specialty hardware?
- Q65.) Whatever is that Hadoop daemon? Explain their roles.
- Q66.) Define “Checkpointing”. What is its benefit?
- Q67.) Name the methods in which Hadoop code can be run.
- Q68.) What is the hoop's map reduction?
- Q69.) How does the Hoodoo map work?
- Q70.) Explain what happens in MapReduce
- Q71.) Explain the Cache distributed in MapReduce structure?
- Q72.) What is the name of Hope's name?
- Q73.) What is JobTracker in Hupa? What does hatio continue to do?
- Q74.) Explain what is heart rate in HDFS?
- Q75.) Explain what the connectors do, and explain that you should use a partner in the mopredos work.
- Q76.) What happens when a data node fails?
- Q77.) What is Special Execution?
- Q78.) What are the basic parameters of the modeler?

- Q79.) What is the MapReduce partition function?
- Q80.) What is the difference between input separation and HDFS block?
- Q81.) What happens in text form?
- Q82.) Do you specify the main configuration parameters that the user should be prompted to work with?
- Q83.) What would you explain to WebDAV at Hadoop?
- Q84.) Explain how JobTracker should schedule a task?
- Q85.) Explain what is Sequencefile?
- Q86.) explain what the conf.setMapper class does?
- Q88.) What is the difference between RDBMS and Hadoop?
- Q89.) Are you familiar with the Hatoyo Gore elements?
- Q90.) What is the name of the hottest?
- Q92.) What is the data storage component used by Hadus?
- Q94.) What is InputSplit in Hupa?
- Q95.) How to Write a Custom Partition for a Hadoop Work?
- Q96.) Can I change the number of producers for a job in the hoate?
- Q97.) What is a visual file in the hottie?
- Q98.) What will happen to the worker?
- Q99.) How is it being indexed in HDFS?
- Q100.) Can search files using wilddo?
- Q101.) Do you want to list three configuration files of hidopo?
- Q102.) Can you verify that the Yaminot uses the JPS command?
- Q103.) 'map' and 'defect' in the hottie?
- Q104.) In the hottest, what file controls do you report into the hoopoe?
- Q105.) To use the network needs of Hadoop?
- Q106.) What is rack awareness?
- Q107.) What is a task tracker in the hatton?

- Q108.) What is Demons running at the edge of the master and slavery?
- Q109.) How can you fix the Hope code?
- Q110.) What is the calculation of storage and nodes?
- Q111.) Do you mean environmental use?
- Q112.) Do you refer to the next step of Mapper or Methus Dogs?
- Q113.) What is the Number of Normal Participation in Hatoyobil?
- Q114.) What is the purpose of recordReader in Hutcho?
- Q115.) How is RDBMS different from HDFS?
- Q116.) What is meant by Big Data and its five V's?
- Q117.) Explain Hadoop with components?
- Q118.) Explain HDFS as well as YARN?
- Q119.) What are Hadoop daemons? Explain their roles.
- Q120.) How is HDFS different from NAS?
- Q121.) Explain Hadoop 1 vs. Hadoop 2.
- Q122.) Explain active/passive "NameNodes"?
- Q123.) What is the reason for adding/removing Hadoop cluster nodes frequently?
- Q124.) Can same files be accessed by two clients in HDFS?
- Q125.) How are DataNode failures managed by NameNode?
- Q126.) How is a down NameNode handled?
- Q127.) Explain a checkpoint?
- Q128.) Explain fault tolerance of HDFS?
- Q129.) Explain commodity hardware in terms of NameNode or DataNode?
- Q130.) Would you use HDFS for large or small data sets?
- Q131.) Explain HDFS block?
- Q132.) Explain jps command?
- Q133.) Explain "speculative execution"?
- Q134.) Can you restart "NameNode"?



- Q135.) How is HDFS Block different from Input Split?
- Q136.) What are three modes for Hadoop running?
- Q137.) Explain MapReduce?
- Q138.) Explain MapReduce configuration parameters?
- Q139.) Can you perform addition in mapper?
- Q140.) Explain RecordReader?
- Q141.) Explain “Distributed Cache”.
- Q142.) Explain communication process of reducer?
- Q143.) What is “MapReduce Partitioner”?
- Q144.) Explain writing process of custom partitioner?
- Q145) Explain “Combiner”?
- Q146.) Explain “SequenceFileInputFormat”?
- Q147.) What is Apache Pig?
- Q148.) Explain Pig Latin data types?
- Q150.) Explain UDF?
- Q151.) Explain “SerDe”?
- Q152.) Can many users use the default “Hive Metastore”?
- Q153.) Name the default location for “Hive” to store table data?
- Q154.) Explain Apache HBase?
- Q155.) Name some components in Apache HBase?
- Q156.) Name Region Server components?
- Q157.) Explain “WAL”?
- Q114.) What is HBase properties?
- Q158.) Explain Apache Spark?
- Q159.) Is it possible to build “Spark” using Hadoop version?
- Q160.) What is RDD.
- Q161.) Explain Apache ZooKeeper?



Q162.) Explain the configuration of “Oozie” job?

Q163.) Explain Rack Awareness?

greentechnologys @ Porur