Hadoop Cluster Setup:

1. How will you add/delete a Node to the existing cluster?

A) Add: Add the host name/Ip address in dfs.hosts/slaves file and refresh the cluster with $hadoop dfsamin -refreshNodes

Delete: Add the hostname/Ip address to dfs.hosts.exclude/remove the entry from slaves file and refresh the cluster with $hadoop dfsamin -refreshNodes

2. What is SSH? What is the use of it In Hadoop?

A) Secure Shell.

3. How will you setup Password-less SSH?

A) search in this site

4. How will you format the HDFS? How frequently it will be done?

A) $hadoop namnode -format.

Note: Format had to be done only once that to during initial cluster setup.

5. How will you manage the Log files generated in Hadoop cluster?

6. Do you know about cron jobs? How will you Setup?

A) In Ubuntu, go to the terminal and type:

$ crontab -e

this will open our personal crontab (cron configuration file), the first line in that file explains it all, In every line we can define one command to run, and the format is quite simple. So the structure is:

minute hour day-of-month month day-of-week command

For all the numbers you can use lists eg, 5,34,55 in the first field will mean run at 5 past 34 past and 55 past what ever hour is defined.

7. What is the role of /etc/hosts file in setting up of HDFS cluster?

A) For hostname to Ip address maping

8. What is dfsadmin command in Hadoop?

9. If one of the data node is failed to start on the cluster how will you come to know? And what are the necessary actions to be taken now ?

A)Via HDFS web UI, we can see no of decommissioned nodes and we need to rebalance the cluster now

10. What is the impact if namenode fails and what are the necessary action items now?

A) Entire hdfs will be down and we need to restart the namenode after copying fsimage and edits from secondaryNN

11. What is Log4j?

A) Logging Framework

12. How do we set logging level for hadoop daemons/commands?

A) In log4j.properties or in hadoop-env.sh file, hadoop.root.logger=INFO,console (WARN,DRFA)

13. Is there any impact on mapreduce jobs if there is no mapred-site.xml file created in HADOOP_HOME/conf directory but all the necessary properties are difined in yarn-site.xml?

A) no

14. How does Hadoop's CLASSPATH plays vital role in starting or stopping in hadoop daemons.

A) Classpath will contain list of directories containing jar files required to start/stop daemons for example HADOOP_HOME/share/hadoop/common/lib contains all the common utility jar files.

15. What is the default logging level in hadoop?

A) hadoop.root.logger=INFO,console.

16. What is the 'hadoop.tmp.dir' configuration parameter default to ?

A) It is user.name. We need a directory that a user can write and also not to interfere with other users. If we didn't include the username, then different users would share the same tmp directory. This can cause authorization problems, if folks' default umask doesn't permit write by others. It can also result in folks stomping on each other, when they're, e.g., playing with HDFS and re-format their filesystem.

17. How do we verify the status and health of the cluster?

A) Either by HDFS Web UI at http://namenode:50070/ or by $ hadoop dfsadmin -report.

18. What is the reason for the frequent exception connection refused in hadoop?

A) If there is no configuration error at client machine or namenode machine, a common cause for this is the Hadoop service isn't running. If there is problem with Check that there isn't an entry for our hostname mapped to 127.0.0.1 or 127.0.1.1 in /etc/hosts.

19. How do we set a configuration property to be unique/constant across the cluster nodes and no slave nodes should override this?

A) We can achive this by defining this property in core/hdfs/mapred/yarn-site.xml file on namenode with final tag as shown below.

<name>mapreduce.task.io.sort.mb</name>

true

20. Does the name-node stay in safe mode till all under-replicated files are fully replicated?

A)No. The name-node waits until all or majority of data-nodes report their blocks. But name-node will stay in safe mode until a specific percentage of blocks of the system is minimally replicated. minimally replicated is not fully replicated.

HDFS Interview Questions and Answers:

1. What is Default replication factor and how will you change it at file level?

2. Why do we need replication factor > 1 in production Hadoop cluster?

3. How will you combine the 4 part-r files of a mapreduce job?

A) Using hadoop fs -getmerge

4. What are the Compression techniques in HDFS and which is the best one and why?

5. How will you view the compressed files via HDFS command?

A) hadoop fs -text

6. What is Secondary Namenode and its Functionalities? why do we need it?

7. What is Backup node and how is it different from Secondary namenode?

8. What is FSimage and editlogs and how they are related?

9. what is default block size in HDFS? and why is it so large?

10. How will you copy a large file of 50GB into HDFS in parllel? A) distcp

11. what is Balancing in HDFS?

12. What is expunge in HDFS ? A) Trash empty

13. What is the default uri for HDFS WEB UI? Can we create files via HDFS WEB UI?

A) namenode:50070. No. It is read only

14. How can we check existence of non zero length file in HDFS commands

A) hadoop fs -test command

15. What is IOUtils in HDFS API and how is it useful?

16. Can we archive files in HDFS? If yes, how can we do that?

A) hadoop archive -archiveName NAME -p <parent path> src dest

17. What is safemode in Hadoop and what are the restrictions during safemode?

18. What is rack awareness in hadoop?

19. Can we come out of safe mode manually, if yes how?

A) $ hadoop dfsadmin -safemode enter/get/leave

20. Why block size in hadoop is maintained as very big compared to traditional block size?

21. What are Sequence files and how are they different from text files?

22. What is the limitation of Sequence files?

A) supports only java, no other API

23. What are Avro files ?

24. Can an avro file created in Java in machine 1 can be read on machine with Ruby API? A) Yes

25. Where does the schema of an Avro file is store if the file is transferred from one host to another?

A) in the same file itself as a header section

26. How do we handle small files in HDFS?

A) merge into sequence/avro file or archive them into har files.

27. What is delegation token in Hadoop and why is it important?

28. What is fsck in Hadoop?

29. Can we append data records to an existing file in HDFS?

A) Yes by command $ hdfs dfs -appendToFile … Appends single src, or multiple srcs from local file system to the destination file system. Also reads input from stdin and appends to destination file system.

30. Can we get count of files in a directory on HDFS via command line?

A) Yes by using command $ hdfs dfs -count hdfs://NN/file1
31. How do we achieve security on Hadoop cluster?

A) With Kerberose

32. Can we create multiple files in HDFS with different block sizes?

Yes. HDFS provides api to specify block size at the time of file creation. Below is the method signature:

public FSDataOutputStream create(Path f, boolean overwrite, int bufferSize, short replication, long blockSize) throws IOException;

33. What is the importance of dfs.namenode.name.dir?

It contains the fsimage file for namenode, it should be configured to write to atleast two filesystems on different physical hosts, namenode and secondary namenode, as if we lose fsimage file we will lose entire HDFS file system and there is no other recovery mechanism if there is no fsimage file available.

34. What is the need for fsck in hadoop?

it can be used to determine the files with missing blocks.

35. Does HDFS block boundaries be between records or across the records?

No, HDFS does not provide record-oriented boundaries, So blocks can end in the middle of a record.

Mapreduce Interview Questions and Answers:

1. What is Speculative execution?

2. What is Distributed Cache?

3. WorkFlow of MapReduce job?

A) map,combiner,reducer,shuffle,partitioner

4. How will you globally sort the output of mapreduce job?

A) totalorder partitioner

5. Difference between map side and reducer side Join?

6. What is Map reduce chaining?

7. How will You pass parameters to mapper or reducer?

8. How will you create custom key and value type's?

9. Sorting based on any column other than Key?

10. How will you create custom input formats?

11. How will you process huge number of small files in MR job?

A) After converting into sequence file/avro file

12. Can we run Reducer without Mapper?

A) Yes in this Identity mapper will be run in the back ground to copy the input to reducer

13. Whether mapper and reducer tasks run in parallel? If no, why see some times as (map 80%,reduce 10%)?

A) No, its due to data copy phase.

14. How will you setup a custom counter to detect bad records in the input?

A) context.getcounter.enumvalue

15. How will you schedule mapreduce Jobs?

A) Through Oozie or Azkaban

16. what is combiner?Tell me one scenario where it is not suitable?

A) for aggregate functions

17. How will you submit mapreduce job through command line?

18. How will you kill a running mapreduce job?

19. For a failed mapreduce job how will trace for the root cause

A) Yarn WEB UI ? logs –> Userlogs ? Application ID container ? Syserr/syslog/

20. What will you do if a mapreduce job failed with Java heap space error message?

A) In HADOOP_CLIENT_OPTS or JAVA_CHILD_OPTS increase Xmx property

21. How many map tasks & reduce tasks will run on each datanode by default

A) 2 map tasks and 1 reduce task

22) What is the minimum RAM capacity needed for this datanode?

As there 3 jvms running for 3 tasks, 1 data node daemon also runs, so, it is needed at least 4 GB RAM, assuming that at least 1GB can be asssigned for each YARN task.

22. What is difference between Mapreduce and YARN?

23. What is Tez framework?

A) An alternative framework for mapreduce, it can be used in Yarn in place of mapreduce

24. What is the difference between Tez and Mapreduce ?

A) Tez is at least 2 times faster than Mapreduce

25. What is input split, input format and record reader in Mapreduce programming?

26. Does Mapreduce support processing of Avro files ? If yes, what is the main classes of the API?

27. How will you process a dataset in JSON format in mapreduce job?

A) JSONObject class can be used to parse the JSON records in the dataset

28. Can we create multi level directory structure (year/month/date) in Mapreduce based on the input data?

A) yes by using multipleoutputs

29. What is the relation between TextOutputFormat and KeyValueTextInputFormat?

A) second one is used to read the files created by first one

30. What is LazyOutpuFormat in Mapreduce and why do we need it?

A) creates output files if data is present

31. How do we prevent file splitting in Mapreduce ?

A) by returning false from isSplittable method on our custom InputFormat Class

32. What is the difference between Writable and WritableComparable interfaces? And what is sufficient for value type in MR job?

A) writable

33. What is the Role of Application Master in running Mapreduce job through YARN?

34. What is Uber task ?

35. What are IdentityMapper & IdentityReducer classes?

36. How do we create jar file with .class files in a directory through command line?

37. What is the default port for YARN Web UI?

A) 8088

38. How can we distribute our application's jars to all of the nodes in the YARN cluster that need it?

39. How do We include native libraries in YARN jobs?

A) by using -Djava.library.path option on the command or else by setting LD_LIBRARY_PATH in .bashrc file.

40. What is the default scheduler inside YARN framework for starting tasks?

A) CapacityScheduler

41. How do we handle record bounderies in Text files or Sequence files in Mapreduce Inputsplits?

In Mapreduce, InputSplit's RecordReader will start and end at a record boundary. In SequenceFiles, every 2k bytes has a 20 bytes sync mark between the records. These sync marks allow the RecordReader to seek to the start of the InputSplit, which contains a file, offset and length and find the first sync mark after the start of the split. The RecordReader continues processing records until it reaches the first sync mark after the end of the split. Text files are handled similarly, using newlines instead of sync marks.

42. Some times mapreduce jobs will fail if we submit the same jobs from a different user? What is the cause and how do we fix these?

A) It might be due to missing of setting mapreduce.jobtracker.system.dir

43. How to change the default location of mapreduce job's intermediate data ?

A) by chaning the value in mapreduce.cluster.local.dir

44. If a map task is failed once during mapreduce job execution will job fail immediately?

A) No it will try restarting the tasks upto max attempts allowed on map/reduce tasks, by default it is 4

Pig Interview Questions and answers:

1. How will load a file into pig?

2. What are the complex data types in pig?

3. What is outer bag?

4. Load an emp table file with columns id, name, deptid, description. Display name and id where deptid="";

5. How will you write custom UDFs?

6. What is the difference between inner bag and outer bag?

7. What is a tuple?

8. What is the difference between FOREACH and FILTER?

9. What is the difference between local mode and mapreduce mode?

10. What is the difference between GROUP BY and JOIN BY in Pig?

11. How many reduce tasks will be run if we specify both GROUP BY and ORDER BY clauses in the same pig script?

12. What is DISTINCT operator?

13. Difference between UNION, JOIN and CROSS ?

14. How do we sort records in descending order in a dataset in Pig? (ORDER DESC/ASC)

15. What is the difference between GROP and COGROUP?

16. What is the difference between STORE and DUMP commands?

17. How will you debug a pig script ? A) set debug on

18. Can we run basic Hadoop fs commands in Grunt shell?

A) yes

19. Can we run Unix shell commands from Grunt shell itself ?

A) yes by using sh command

20. Can we submit pig scripts in batch mode from grunt shell?

A) yes by using run/exec command

21. What is the difference between run and exec commands in grunt shell?

A) Run will execute the pig script in the same grunt shell but exec will submit in a new grunt shell

22. What are diagnostic operators in Pig?

23. What is the difference between EXPLAIN, ILLUSTRATE and DESCRIBE?

24. How do we access a custome UDF function created in Pig?

A) by using REGISTER and DEFINE functions it will be available in pig session

25. What is DIFF function in pig?

26. Can we do random sampling from a large dataset in pig?

A) SAMPLE command

27. How can we divide records of a single dataset into multiple datasets by using any criteria like country wise?

A) using SPLIT command

28. What is the difference between COUNT and COUNT_START functions in pig?

A) COUNT_START includes null values also in counting whereas COUNT will not

29. What are PigStorage & HBaseStorage ?

30. What is the use of LIMIT in pig?

31. What is the difference between Mapreduce and Pig and can we use Pig in all scenarios where we can write MR jobs?

A) No

Hive Interview Questions and Answers:

1. Does hive support record level operations?

2. In hive table can we change string DT to Int DT?

3. Can we rename a Table in Hive? if Yes, How?

4. What is metastore? how will you start the service?

5. WHat is Serde in Hive?Example?

6. Difference between Hive and Hbase?

7. How to print column name of a table in hive query result?

8. How will you know whether a table is external or managed?(desc extended)

9. What is Hive thrift server?

10. What is the difference between local metastore and embedded metastore?

11. How do we load data into Hive table with SequenceFile format from text file on local file system.

12. What is HCatalog?

13. How is HCatalog is different from Hive?

14. What is WebHCat?

15. How do we import XML data into Hive?

16. How do we import CSV data into Hive?

17. How do we import JSON data into Hive?

18. What are dynamic partitions?

19. Can a Hive table contain data in more than one format?

20. How do I import Avro data into Hive?

21. Does Hive have an ODBC driver?

A) Yes cloudera provides ODBC drivers for Hiveserver

22. Is HiveQL case sensitive? A) no

23. Does Hive support Unicode?

A) Yes we can use Unicode string on data/comments, but cannot use for database/table/column name.

24. Can a Hive table contain data in more than one format?

25. Is it possible to set the data format on a per-partition basis?

26. What are dynamic partitions?

27. Does Hive have a JDBC Driver?

A) Yes, The driver is 'org.apache.hadoop.hive.jdbc.HiveDriver'. It supports two modes: a local mode and a remote one. In the remote mode it connects to the hive server through its Thrift API. The JDBC url to use should be of the form: 'jdbc:hive://hostname:port/databasename'. In the local mode Hive is embedded. The JDBC url to use should be 'jdbc:hive://'.

28. How can we import fixed-width data into Hive?

29. How can we import ASCII logfiles (HTTP, etc) into Hive?

30. When running a JOIN query, what is the idea to solve out-of-memory errors.

A)This is usually caused by the order of JOIN tables. Instead of "FROM tableA a JOIN tableB b ON …", try "FROM tableB b JOIN tableA a ON …". NOTE that if we are using LEFT OUTER JOIN, we might want to change to RIGHT OUTER JOIN. This trick usually solve the problem – the rule of thumb is, always put the table with a lot of rows having the same value in the join key on the rightmost side of the JOIN.

31. How many times Tez engine runs faster than MR engine in Hive?

32. How much time each Tez session will be active?

Hbase Interview Questions and answers:

1. What are the Catalog tables in Hbase?

2. What is Zookeeper role in hbase architecture?

3. How will you drop a table in Hbase?

4. Do you know Hive on hbase? how will you achive it?(Hbasestorage handler)..If we delete a table from hive will it effect on hbase table?

A) yes

5. How will you load bulk data of 50GB file into Hbase table?

6. Limitations of Hbase?(no support for sql syntax, indexing, joins,..)

7. Difference between Hbase and Hdfs?

8. How do we integrate HBase and Hive?

9. How can we add/remove a node to HBase cluster?

A)By Adding/removing an entry in HBASE_CONF_DIR/regionservers file

10. Can we safely move the hbase rootdir in hdfs?

A) Yes. HBase must be down for the move. After the move, update the hbase-site.xml across the cluster and restart.

11. Can we safely move the master from node A to node B?

A) Yes. HBase must be down for the move. After the move, update the hbase-site.xml across the cluster and restart.

12. How do we fix OutOfMemoryExceptions in hbase?

A) Hbase uses a default of 1 GB heap size. By increasing this at HBASE_HEAPSIZE environment variable in ${HBASE_HOME}/conf/hbase-env.sh we can solve these error messages.

13. How can we change logging level in HBase?

A) In log4j.properties file we can set logging level as DEBUG–
log4j.logger.org.apache.hadoop.hbase=DEBUG — and restart our cluster or in hbase-env.sh file.

14. What ports does HBase use?

A) hbase runs the master and its informational http server at 60000 and 60010 respectively and regionservers at 60020 and their informational http server at 60030.

15. Some times HBase is ignoring HDFS client configuration such as dfs.replication. what is the cause?

A) If we made HDFS client configuration on our hadoop cluster, HBase will not see this configuration unless:
We Add a pointer to HADOOP_CONF_DIR to CLASSPATH in hbase-env.sh or symlink your hadoop-site.xml from the hbase conf directory.
Add a copy of hadoop-site.xml to ${HBASE_HOME}/conf, or If only a small set of HDFS client configurations, add them to hbase-site.xml
The first option is the better of the three since it avoids duplication.

16. What is the maximum recommended cell size?

A) A rough rule of thumb, with little empirical validation, is to keep the data in HDFS and store pointers to the data in HBase if you expect the cell size to be consistently above 10 MB. If you do expect large cell values and you still plan to use HBase for the storage of cell contents, you'll want to increase the block size and the maximum region size for the table to keep the index size reasonable and the split frequency acceptable.

17. Why can't I iterate through the rows of a table in reverse order?

A) Because of the way HFile works: for efficiency, column values are put on disk with the length of the value written first and then the bytes of the actual value written second. To navigate through these values in reverse order, these length values would need to be stored twice (at the end as well) or in a side file. A robust secondary index implementation is the likely solution here to ensure the primary use case remains fast.

18. What is phoenix?

A) phoenix is an sql layer on hbase

19. How fast is Phoenix? Why is it so fast?

A) Phoenix is fast. Full table scan of 100M rows usually completes in 20 seconds (narrow table on a medium sized cluster). This time come down to few milliseconds if query contains filter on key columns. For filters on non-key columns or non-leading key columns, you can add index on these columns which leads to performance equivalent to filtering on key column by making copy of table with indexed column(s) part of key.

20) Why is Phoenix fast even when doing full scan?

A) Phoenix chunks up your query using the region boundaries and runs them in parallel on the client using a configurable number of threads
The aggregation will be done in a coprocessor on the server-side, collapsing the amount of data that gets returned back to the client rather than returning it all.

Sqoop Interview Questions and answers:

1. How will you get data from RDBMS into HDFS?

2. Can we store mysql table data as sequence file in hdfs via sqoop?

3. Does sqoop support compression techniques to store data in HDFS?

4. Can we load all the tables in a database into hdfs in a single shot?

A) import-all-tables

5. Can we copy a subset of data from a table in RDBMS into HDFS?(based on some criteria)

A) Using –where "country='us'" condition in import command

6. How many reduce tasks will be run by default for a sqoop import command?How many mappers?

A) 0 , 4

7. If we get java heap space error and we have already given the maximum memory, what is the possible solution?

A) increase mappers by -m 100

8. What is the default port for connecting to MySQL server?

A) 3036

9. How can we resolve a Communications Link Failure when connecting to MySQL?

Verify that we can connect to the database from the node where we are running Sqoop:
$ mysql –host= –database=test –user= –password= Add the network port for the server to your my.cnf file. Set up a user account to connect via Sqoop. Grant permissions to the user to access the database over the network:
Log into MySQL as root mysql -u root -p
Issue the following command: mysql> grant all privileges on *.* to 'user'@'%' identified by 'testpassword'
mysql> grant all privileges on *.* to 'user'@" identified by 'testpassword'

10. Can we provide SQL queries in SQOOP Import command?

Flume Interview Questions and answers:

1. Can we load data directly into Hbase?

A) yes

2. How will you create directories in HDFS based on the timestamp present in input file?

A) hdfs.path=/user/%y-%m-%d/%H%M%S)(format escape sequences)

3. What will happen if no timestamps are present in input file?

it will throw an exeception, to slove this hdfs.useLocalTimeStam=true...

4. Work flow of flume?

5. What are the channel types in Flume?(Memory,JDBC,File channel)Which one is faster – memory?

6. How will you start a flume agent from Command line?

7. What are interceptors in flume?

8. We are getting a NumberFormatException when using format escape sequences for date & time(%Y %M %D etc..) in HDFS sink. How can we solve this exception?

To use data escape sequences in Flume, there should be timestamp present in header of the source record. If there is no timestamp in the source file, we can solve this exception by two ways
i) By adding Timestamp interceptor in source as shown below
a1.sources.tail.interceptors = ts
a1.sources.tail.interceptors.ts.type =
org.apache.flume.interceptor.TimestampInterceptor$Builder
ii) Or by adding Use local timestamp=true parameter in configuration properties of agent for HDFS sink.

9. What is the bridge mechanism used for Multi-hop agent setup in Flume?

A) Avro RPC

9. Which is the reliable channel to make sure there is no data loss (JDBC, File, Memory)?

A) file channel is reliable

10. What is Fan out flow in Flume?

11. What are the event serializers available in Flume?

A) Text, Avro

12. How do we collect records in JSON format directly through Flume?

A) by usring JSONHandler

13. What is the difference between FileSink and File Roll Sink?

14. Difference between ASynchHbase Sink and HBase sink types ?

15. If We need to test the functionalities of a custom source and channel and we do not need any sink can we setup this kind of agent?

A) Yes, with sink type as null

16. Can we perform realtime analysis on the data collected by Flume directly ? if yes how?

A) yes by using MorphlineSolrSink we can extract data from Flume events, transform it, and load it in near-real-time into Apache Solr servers, which in turn serve queries to end users or search applications.

17. If we need to get speed of memory channel and data reliabilty of file channel in a single agent channel, then how can we achieve this?

A) Use SpillableMemoryChannel for this purpose

18. What are multiplexing selectors in flume?

19. What are replication selectiors in flume?

20. What is the use of HostInterceptor in flume?

21. What is the advantage of UUIDInterceptor in flume?

22. In defining type of sources or sinks in flume is it mandatory to provide the full class name?

A) No, we can also provide the alias names. For example, we use hdfs as sink.type in place of org.apache.flume.sink.hdfs.HDFSEventSink

Splunk Interview Questions and answers:

1. What is Splunk and what is hunk?

2. How do we connect to HDFS in hunk?

3. Is there any connector for Hive server directly to load Hive tables into Hunk?

4. What is HiveSplitgenerator in Hive provider?

5. Do we need to keep Hive thrift server running and Hive metastore services running to retrieve hive tables into hunk? yes

6. Can we create dashboards in hunk with visualization charts embedded in it?

7. Deos Hunk support reading of compressed files (.gz, .bz2) files on Hadoop?

8. Does hunk support reading of snappy compressed file on hadoop?

9. Where do we can look for the error messages or exceptions in search query in hunk? (search.log file under dispatcher folder in Hunk distribution)

10. What is the default port for accessing Hunk web UI? (8000)

Tableau Interview Questions and answers :

1. Can we use Tableau on Linux server ? (no, supports only windows and mac)

2. What is the difference between Hunk and Tableau

3. How do we connect to Hiveserver from Tableau?

4. How can we connect to MySQL from tableau?

5. Can we perform data blending of two different sources in tableau?

6. Do we need to write queries to perform joins or filters in tableau?

7. Does Tableau fire any mapreduce jobs in the backend to pull data from hive?

Oozie & Azkaban Interview Questions and answers:

1. What is the Job scheduler you use in your production cluster?

2. Does Oozie support time based triggering of jobs ? (yes)

3. Does Azkaban support time based triggering of jobs ? (yes)

4. Does Oozie support data based triggering of jobs ? (yes)

5. Does Azkaban support time based triggering of jobs ? (yes)

6. Can we define dependencies between jobs in Akaban flows? (yes)

7. What is the difference between Oozie and Azkaban?

8. How do we create properties files in Azkaban?

9. How do we create properties files in Oozie?

Unix Interview Questions and answers:

1. How do you know what are the processes running in Unix?

$ ps – lists all the unix system processes

$ jps – Lists all the java processes

$ jobs – Lists all the processes that were suspended and running in the

background. Because the jobs command is a foreground process, it cannot show us active foreground processes.

2. How will you stop a process forcibly in Unix?

Use the below command to kill/stop a process forcibly.

$ kill -9 processid;

Here option -9 denotes force killing.

3. Will the below commands result in same output ?

TEST="hello world" ;

$ echo $TEST;

$ echo TEST;

Ans) No. first command will print "hello world" on console and second one will print "TEST" on console.

4) How can we define constants in Unix shell scripting?

Ans) We can achieve this with the help of readonly variables in Unix shell scripting.

For example, consider the following commands:

$ TEST1="hello"

$ readonly TEST1

$ echo $TEST1

hello

$ TEST1="world"

The last command results in an error message:

/bin/sh: TEST1: This variable is read only.

5) Can we unset variables in Unix ?

Yes, we can release the variable names by using unset command.

For example,

$ unset TEST;

will release the variable TEST and it no longer references "hello world" string. But we cannot use the unset command to unset variables that are marked readonly.For example,

$ unset TEST1;

will result in an error message.

6) What are Environment Variables in Unix?

An environment variable is a variable that is available to any child process of the shell. We will make a variable environmental by using export command.

Syntax for declaring environment variables: $ name=value ; export name;

7) What are Shell Variables in Unix?

A shell variable is a special variable that is set by the shell and is required by the shell in order to function correctly. Some of these variables are environment variables whereas others are local variables.

These are the variables that the shell sets during initialization and uses internally. Example are:

PWD Indicates the current working directory as set by the cd command.

UID Expands to the numeric user ID of the current user, initialized at shell startup.

PATH Indicates search path for commands. It is a colon-separated list of directories in which the shell

looks for commands. A common value is

HOME Indicates the home directory of the current user: the default argument for the cd built-in command.

8) What does $@ represent in Unix? (All arguments of command)

9) What is $? in unix(it is the status of last executed command)

10) What is sed? and why do use it? (It is stream editor, it can be used for replacing set of characters with other set)