WEEK 7 QUIZ SOLUTION

1. Which of these is true- “Each row in HBase table is?
    a. Rigid at both column level & column family level
    b. *Rigid at column family level & flexible at column level
    c. Rigid at column level & flexible at column family level
    d. Flexible at both column level and column family level

    Explanation: HBase does not need to fit into a rigid schema like an RDBMS and its rigid at column family level but columns are flexible. You can add as many columns you need in column family dynamically. Column families cannot be added dynamically.

2. Every new insert, in HBase is written to HDFS instantly. True/False
    a. TRUE
    b. *FALSE

    Explanation: HBase new insert goes to WAL and MemStore first. The MemStore is a write buffer where it stores data in memory before a permanent write.

3. HFile is stored in ____, HLog is stored in _____?
    a. Memory, HDFS
    b. HDFS, Memory
    c. *HDFS, HDFS
    d. Memory, Memory

    Explanation: MemStore once filled flushed periodically in HFile. HFile is nothing but HDFS repository.

    HBase maintains the in-memory log file called HLog. This file contains the updates happening in tables. This cache is flushed periodically to HDFS

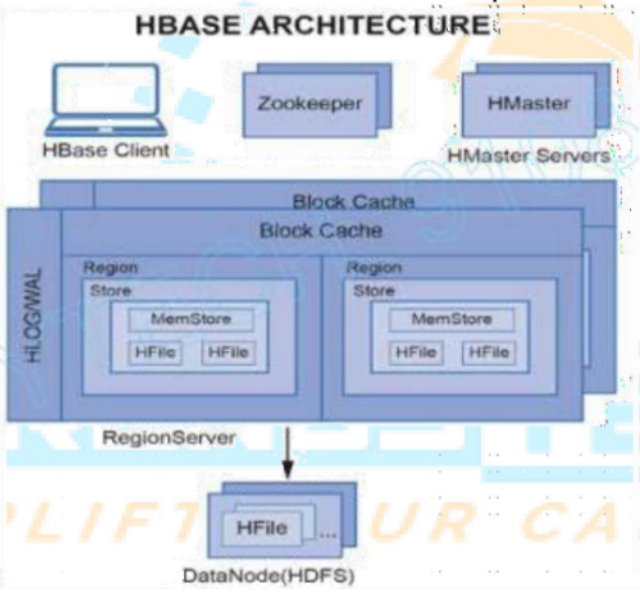
4. Suppose in HBase we have 3 Region servers, 3 regions and 2 column families. How many MemStore are possible?
    a. 1
    b. 3
    c. 2

d. *6

Explanation: RegionServer holds Region and Region holds store. One store represents one Column family and every CF has one MemStore. Hence region * CF = MemStore for e.g., 3*2=6

5. In above case, how many total WAL and Block Cache are possible?
   a. *3,3
   b. 1, 1
   c. can't be fixed
   d. 6, 6

   Explanation: WAL and BlockCache are present for every RegionServer.



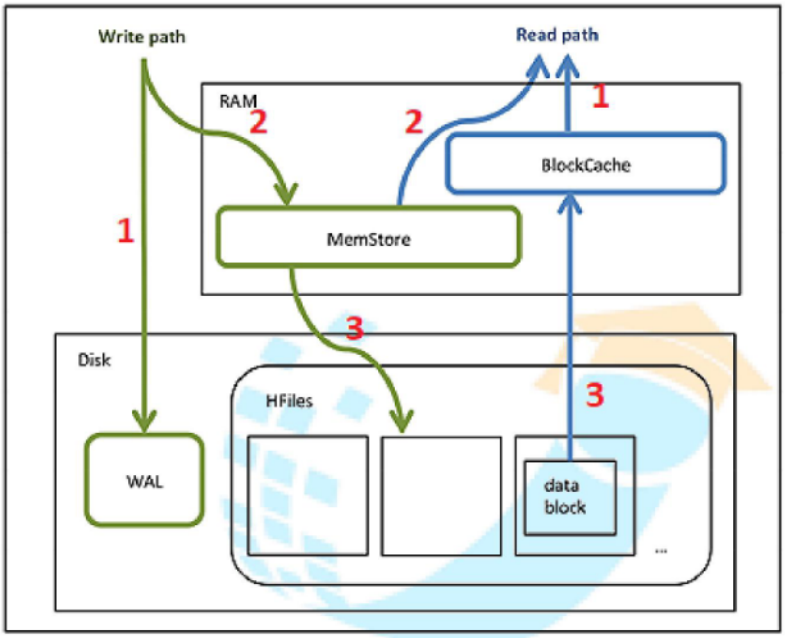
6. Identify the correct sequence for HBase inserts?
   a. MemStore -> HFile -> HLog
   b. HFile -> MemStore -> HLog
   c. *HLog -> MemStore -> HFile
   d. HLog -> HFile -> MemStore

   Explanation: Write process first writes to HLog then to MemStore and then HFile

7. Which of the HBase component can offer fastest reads?
   a. MemStore

b. WAL
c. HMaster
d. *Block Cache

Explanation: Nearest location is Block cache and hence fastest



8. Which of these is not true ,multiple can be chosen.
   a. *Zookeeper helps in Load Balancing of region servers
   b. *Zookeeper holds the MetaTable
   c. *hMaster is Master and Regions are Slaves
   d. *WAL - In memory read/write buffer

   Explanation: true statements are

   a. HMaster helps in Load Balancing of region servers
   b. RegionServer holds the MetaTable and its location is stored in Zookeeper
   c. HMaster is Master and RegionServer are Slaves
   d. WAL - In memory write buffer

9. Region Server crashes are handled by ____, while data loss due to Region Server crashes are recovered by _____?

a. Zookeeper, WAL
b. Zookeeper, MetaTable
c. *hMaster, WAL
d. WAL, HFile

Explanation: Region Server crashes are handled by HMaster, while data loss due to Region Server crashes are recovered by WAL

10. A Region Server might contain-choose most appropriate one:
    a. Region
    b. Region, MemStore
    c. *Region, BlockCache, WAL, MetaTable
    d. Region, WAL, MemStore

    Explanation: RegionServer consists of **Region, BlockCache, WAL, MetaTable**

11. Point of interaction between a client and HBase,for very first time is ?
    a. HMaster
    b. RegionServer
    c. *Zookeeper
    d. None of the above

    Explanation: The client first contacts the zookeeper to fetch the location of the MetaTable.

12. HBase cluster overall health is monitored by
    a. RegionServer
    b. HMaster
    c. HLogs
    d. *Zookeeper

    Explanation: Zookeeper monitors the health of HBase. RegionServer keeps sending the heartbeat to zookeeper.

13. MetaTable can be cached for quick retrieval in?
    a. Zookeeper

b. *Client Machine
c. HMaster
d. MemStore

Explanation: The client first contacts the zookeeper to fetch the location of the MetaTable(if it does not exists). Normally it gets cached on client machine

14. Correct HBase Read Sequence Preference for quicker reads?
    a. HFile -> MemStore -> BlockCache
    b. MemStore -> HFile -> BlockCache
    c. *BlockCache -> MemStore -> HFile
    d. MemStore -> BlockCache -> HFile

    Explanation: During Read Process, it will first try to find the data in BlockCache, if not found then it searches in MemStore(write buffer) and then finally looks for the data in HFILE.

15. Pick the odd one out:
    a. HBase is a database built on top of the HDFS.
    b. HBase provides fast lookups for larger tables.
    c. It provides low latency access to single rows from billions of records (Random access).
    d. *It provides only sequential access of data.

    Explanation: It provides random access of data

16. As per CAP theorem, a system can satisfy at most 2 of 3 guarantees i.e. Consistency, Availability and Partition Tolerance. Which of the following guarantees cannot be sacrificed in a distributed system?
    a. Consistency
    b. Availability
    c. *Partition Tolerance
    d. None of the above.

    Explanation: Partition tolerance cannot be sacrificed in a distributed system

WEEK 7 QUIZ SOLUTION

17. Which of the following statement is true about HBase?
    a. HBase is governed by its schema, which describes the whole structure of tables.
    b. *HBase is schema-less, it doesn't have the concept of fixed columns schema; defines only column families.
    c. HBase is thin and built for small tables. Hard to scale.
    d. HBase will have normalized data.

    Explanation: B is true , all others are false.

18. What would be the likely outcome of the following shell command?
    delete 'employee', '1', 'personal details:city', 1423524448375
    a. It will delete the cell from the EMPLOYEE table immediately
    b. *The cell will become invisible and will remain in the server for some time
    c. Will generate an error
    d. The command will execute but the cell will be visible in the EMPLOYEE table.

    Explanation: When you delete a row, data is not deleted, instead its marked as tombstone marker, making the data effectively invisible. HBase periodically removes tombstone marker during major compactions.

19. Whenever a client wants to insert new data into an HBase table, it is first written to
    a. *Write Ahead Log
    b. MemStore
    c. BlockCache
    d. HFile

    Explanation: During write, data first goes to WAL.

20. Suppose the EMPLOYEE table is no longer required and you want to delete this table. Which of the following HBase shell command/s will be used to delete this table? 1.delete 2. discard 3. drop 4. disable
    a. only 1
    b. only 2
    c. 2 and 3
    d. *3 and 4

    Explanation: drop and disable is used to delete the table.