Quiz 6:

01 Which of the scenarios below provide a strong case to choose HBase over a traditional RDMS (assuming data can be stored in rows/columns)?
o Source data and business requirements may frequently change causing frequent addition / deletion of rows in the database
o Schema of the database is well known in advance and it is unlikely to change over time
X Source data and business requirements may frequently change causing frequent addition / deletion of table attributes in the database
o All of the above

02 Which of the following requirements is considered one of the motivations for the birth of NoSQL technologies?
o Need for strong consistency to support transactions in web applications
X Need for an (relatively) easy to scale solution for data storage
o Need for highly available data stores with rigid schemas
o Need for highly available data stores with strong support for online transaction processing

03 Which of the following statements is true when comparing HBase vs. HDFS?
X HDFS has poor performance (w.r.t. HBase) when random access to data is needed
o HBase does great in batch processing (w.r.t. HDFS), specially for operations involving file scans (in file systems)
o HBase is built on top of HDFS and therefore can only support operations for file scans and record-level insertion.
o None of the above

04 Why do column-based NoSQL technologies typically store column families together on disk?
o Mainly because is semantically convenient (columns in a column family are related to each other) when writing queries
o Mainly because columns in a column family belong to the same HFile
o Mainly because HTables are stored on separate files
X Mainly for performance reason (for accessing data from disk)

05 A BASE data sharing system provides the necessary mechanisms to guarantee that clients will get the latest version of data at any time T.
o True
X False