**1. What is NoSQL data base?**

Ans=> A NoSQL (originally referring to "non SQL" or "non relational")database provides a mechanism for storage and retrieval of data that is modeled in means other than the tabular relations used in relational databases. NoSQL databases are increasingly used in big data and real-time web applications.

NoSQL systems are also sometimes called "Not only SQL" to emphasize that they may support SQL-like query languages.

NoSQL databases are highly scalable and flexible database management systems which allow you to store and process unstructured as well as semi-structured data which is not possible through RDBMS tools.

**2.How does data get stored in NoSQl database?**

Ans:

There are basically 4 ways to store data in NOSQL database.

a) Key-value store

key-value stores are the simplest NoSQL data stores to use. The client can either get the value for the key, assign a value for a key or delete a key from the data store.

b) Document store

Document store NoSQL databases are similar to key-value databases in that there’s a key and a value. Data is stored as a value. Its associated key is the unique identifier for that value. The difference is that, in a document database, the value contains structured or semi-structured data. This structured/semi-structured value is referred to as a document and can be in XML, JSON or BSON format.

c) Column store

In column-oriented NoSQL databases, data is stored in cells grouped in columns of data rather than as rows of data. Columns are logically grouped into column families. Column families can contain a virtually unlimited number of columns that can be created at runtime or while defining the schema. Read and write is done using columns rather than rows. Column families are groups of similar data that is usually accessed together.

d) Graph base

Graph databases are basically built upon the Entity – Attribute – Value model. Entities are also known as nodes, which have properties. It is a very flexible way to describe how data relates to other data. Nodes store data about each entity in the database, relationships describe a relationship between nodes, and a property is simply the node on the opposite end of the relationship. Whereas a traditional database stores a description of each possible relationship in foreign key fields or junction tables, graph databases allow for virtually any relationship to be defined on-the-fly.

**3.What is a column family in HBase?**

In the HBase data model columns are grouped into column families, which must be defined up front during table creation. Column families are stored together on disk, which is why HBase is referred to as a column-oriented data store.

For example, the columns courses:history and courses:math are both members of the courses column family. The colon character (:) delimits the column family from the .

**4.How many maximum number of columns can be added to HBase table?**

There is no limit on number of column families in HBase, in theory.

HBase currently does not do well with anything above two or three column families so keep the number of column families in your schema low. Currently, flushing and compactions are done on a per Region basis so if one column family is carrying the bulk of the data bringing on flushes, the adjacent families will also be flushed though the amount of data they carry is small.

**Q5 Why columns are not defined at the time of table creation in HBase?**

Ans:

Columns in Apache HBase are grouped into column families. Column families must be declared up front at schema definition time whereas columns do not need to be defined at schema time but can be conjured on the fly while the table is up an running.

Q**6 How does data get managed in HBase?**

Ans:HBase is a column-oriented database that’s an open-source implementation of Google’s Big Table storage architecture.

HBase Data Model Terminology

TableAn HBase table consists of multiple rows.

Row :A row in HBase consists of a row key and one or more columns with values associated with them. Rows are sorted alphabetically by the row key as they are stored.

Column::A column in HBase consists of a column family and a column qualifier, which are delimited by a : (colon) character.

Column Family :Column families physically colocate a set of columns and their values, often for performance reasons.

Column Qualifier:A column qualifier is added to a column family to provide the index for a given piece of data. Given a column family content, a column qualifier might be content:html, and another might be content:pdf. Though column families are fixed at table creation, column qualifiers are mutable and may differ greatly between rows.

Cell:A cell is a combination of row, column family, and column qualifier, and contains a value and a timestamp, which represents the value’s version.

**Q7. What happens internally when new data gets inserted into HBase table?**

Ans: Every time you perform an insert operation on a cell, HBase implicitly stores a new version. Creating, modifying, and deleting a cell are all treated identically; they’re all new versions.

When a cell exceeds the maximum number of versions, the extra records are dropped during the next major compaction. Instead of deleting an entire cell, you can operate on a specific version or versions within that cell. Values within a cell are versioned. Versions are identified by their timestamp, a long. When a version isn’t specified, the current timestamp is used as the basis for the operation. The number of cell value versions retained by HBase is configured via the column family. The default number of cell versions is three.