Q 1 Unlike Relational Databases, NoSQL databases are known to be *Scalable* systems that are *Schemaless*.

a) Discuss the challenge of scaling a relational DBMS. **[6 marks]**

Scaling Transaction is difficult, as transaction may update data on multiple machines. and to apply ACID, transaction should be committed using 2PCP

b) Indicate two different approaches that NoSQL systems take to overcome the above challenge. **[6 marks]**

- **Limiting transaction support**
  - **Allow only 1 transaction to be done**
- **Limit query Capabilities**
  - **Prevent Join**
- **Provide eventual consistency**
  - **Instead of strong consistency**

c) Discuss, in the light of Key-value stores, the advantages and disadvantages of a schemaless database design. **[6 marks]**

ADV : Simple & Flexiable

DIS : Designing of key value, then discuss how it's constructed

d) In Document data stores, explain how Keys can play an important role in implementing scalable architectures. **[7 marks]**

- Shard key exist in all documents used to separate them
- Used to specify values which is gonna be grouped to different shards
- Enable many docs to scale to meet demands of apps with heavy loads

Q 2 NoSQL databases can easily be distributed across a cluster of computers.

a) Discuss the two commonly used methods for data distribution in NoSQL Systems.

[7 marks]

- **Sharding**

  Distribute different data on different servers, each node share the same schema

  - Improve read & write and increase fault tolerance

- **Replication**

  Copy data set across multiple servers, so it can be found at any server locally

  - Provide scalability / Availability / Fault tolerance due to replication of data
  - Improve read performance

b) Specify and discuss the two distribution architectures that are used in Column Family Databases. **[7 marks]**

→ CoSandra

- Master-Slave
- Peer to peer