

Lesson 7

Shared Nothing Architecture and Distribution Models

Shared nothing (SN)

- A cluster architecture
- A node does not share data with any other node
- Big Data store consists of SN architecture
- Big Data store, therefore, easily partitions into shards

SN Clusters Partitioning

- A partition processes the different queries on data of the different users at each node independently
- Thus, data processes run in parallel at the nodes
- A node maintains a copy of running-process data

Partitioning

- Data of different data stores partition among the number of nodes (assigning different computers to deal with different users or queries)
- Processing may require every node to maintain its own copy of the application's data, using a coordination protocol.

Partitioning and Processing

- Examples: Hadoop, Flink and Spark.

SN Architecture Characteristics

- Independence: Each node with no memory sharing; computational self-sufficiency
- Self-Healing: A link failure causes creation of another link
- Each node functioning as a shard:
- No network contention

Distribution Models

- Single Server Model
- Master Slave Distribution (MSD) Model
- Peer-to-Peer distribution (PPD) model

Figure 3.9 (a) Single server model (b) Shards distributed on four servers in a cluster

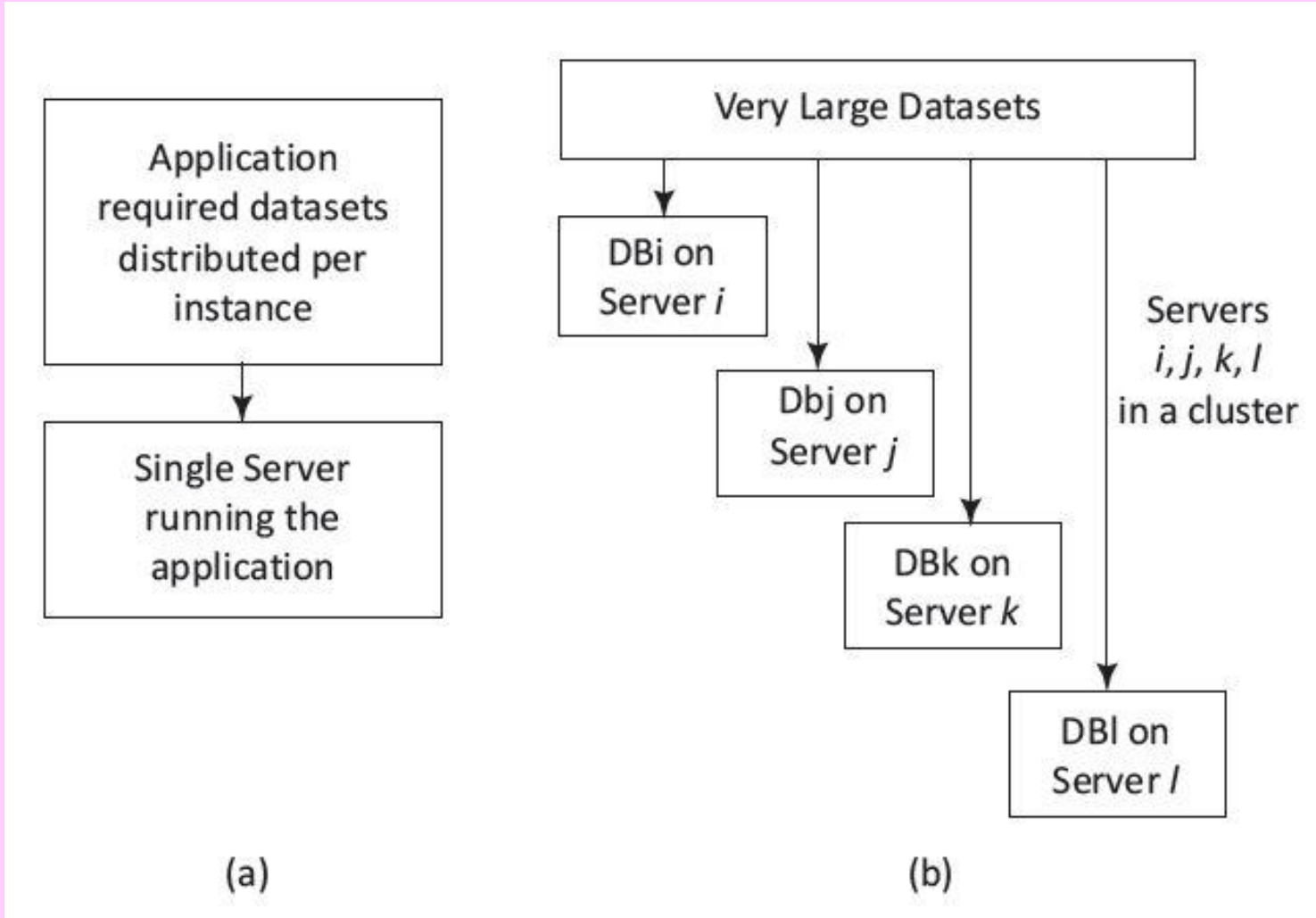


Figure 3.10 Master-slave distribution model. Mongo is a client and mongod is the server

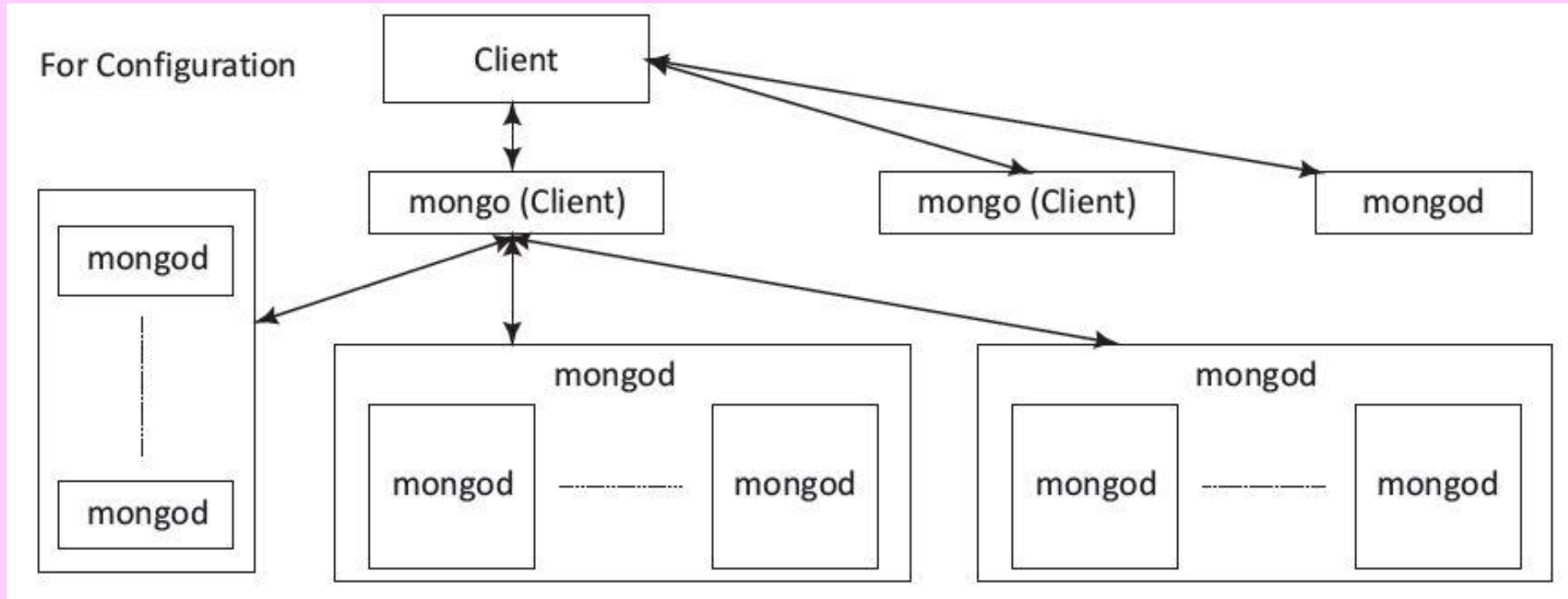


Figure 3.11 PPD Model: Shards replicating on the nodes, which does read and write operations both

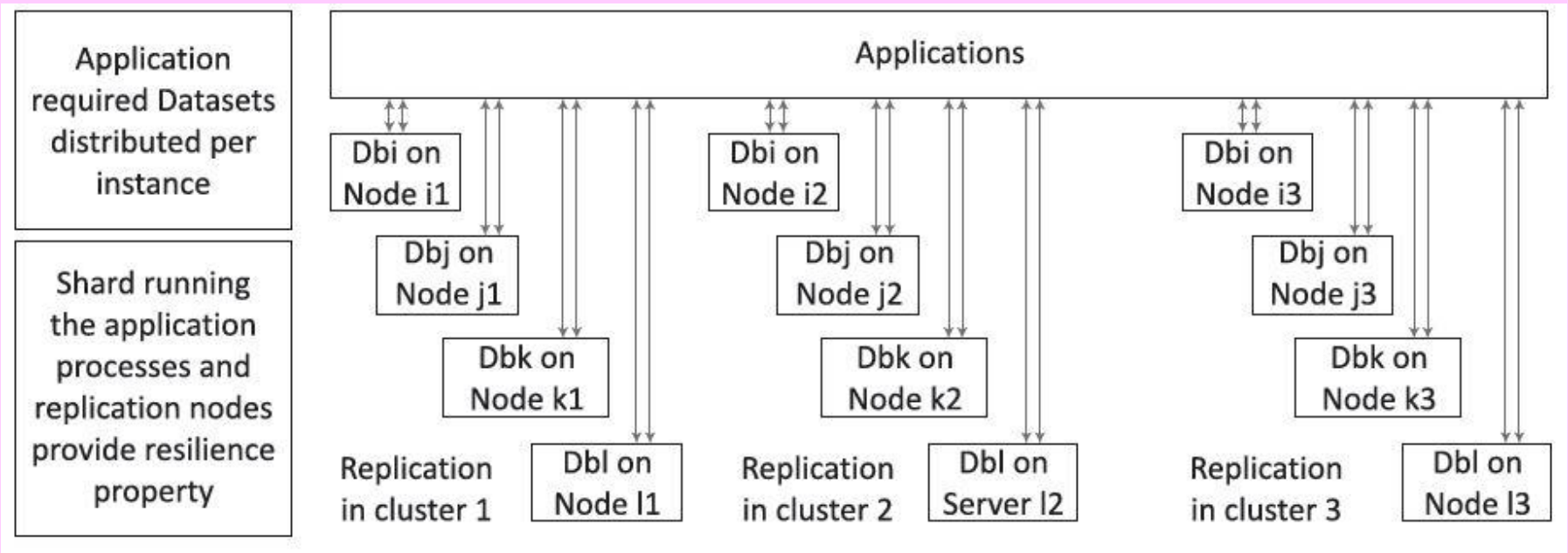
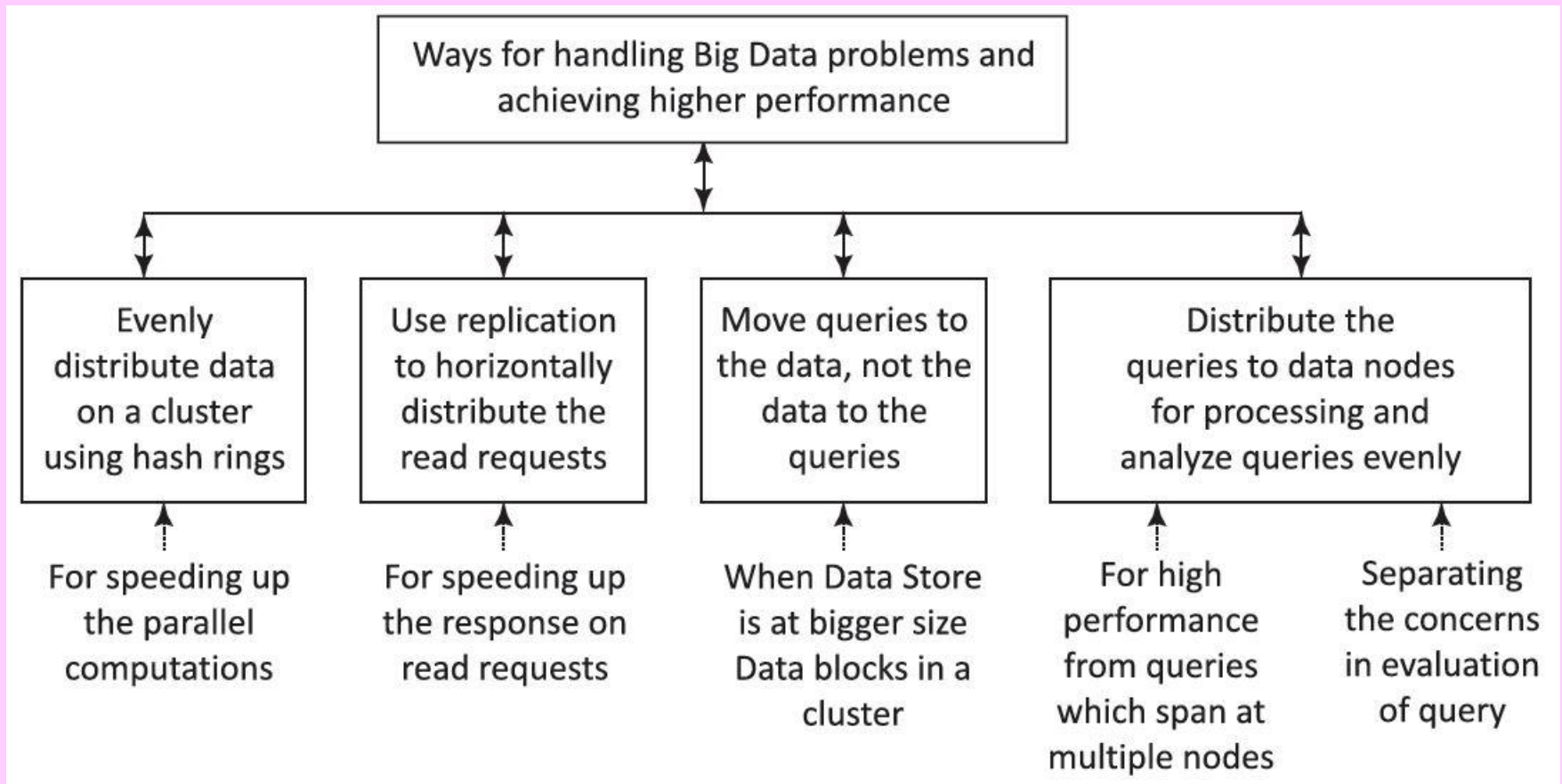


Figure 3.12 Four ways for handling big data problems



Summary

We learnt:

- Shared Nothing Architecture
- Distribution Models
- Single Server
- Master Slave
- Peer to Peer distribution model
- Four ways of handling Big Data Problems

End of Lesson 7 on
**Compare NoSQL data stores with
SQL databases**