|  |  |
| --- | --- |
|  | Apache Cassandra |
| Basics/features | <https://en.wikipedia.org/wiki/Apache_Cassandra>: Open-source, distributed NoSQL database management system |
| Motivation |  |
| Use cases | <https://cassandra.apache.org/>: For apps where data loss is unaffordable |
| Flexibility/scalability | <https://cassandra.apache.org/>: Linear scalability (add more nodes), synchronous/asynchronous replication |
| Failure/fault tolerance | <https://cassandra.apache.org/>: Data replicated to nodes, across data centers. Nodes that fail get replaced fast. |
| Limitations/bottlenecks | <https://cassandra.apache.org/>: Decentralized, no single place of failure, network bottleneck |
| Message consumption/delivery/ordering | <https://medium.com/@anirudhkanabar/clustering-order-in-cassandra-how-to-achieve-ordering-of-data-in-cassandra-f52e8a73d5d5>: Clustering order feature allows queries to be sorted according to their given parameter |
| Storage/retention | <https://www.quora.com/How-much-data-can-a-single-node-in-a-Cassandra-cluster-handle-Are-there-configuration-settings-to-set-the-data-size-limit-the-data-size-on-the-node>/https://stackoverflow.com/questions/4775388/how-much-data-per-node-in-cassandra-cluster: Under 1 TB per node  <https://stackoverflow.com/questions/7190573/cassandra-node-limitations>: A row must fit 1 node. |
| Data transmission protocol/accuracy | <https://stackoverflow.com/questions/28127182/data-transmission-between-nodes-and-client-cassandra>: Gossip  <https://en.wikipedia.org/wiki/Gossip_protocol>: Gossip allows data to spread to all group members. |
| Implementation | <https://www.youtube.com/watch?v=lE8LI5BvDGA>   * Similar to MySQL * Create a keyspace and use it. * Create a table and run queries like INSERT and SELECT. * Can be used in a terminal, so can be programmatically achieved using bash files or Python OS commands |
| Real-world scenarios | <https://cassandra.apache.org/>: Activation, Apple, Best Buy, eBay, Instagram, Netflix, Uber, Walmart, etc. |
| Qualitative/quantitative comparisons w/ other technologies | Vs. MySQL:   * <https://www.geeksforgeeks.org/difference-between-cassandra-and-mysql/#:~:text=Cassandra%20is%20a%20NoSQL%20type,is%20a%20RDBMS%20type%20database.&text=Read%20performance%20is%20highly%20efficient,from%20multiple%20tables%20using%20JOIN>. * <https://stackoverflow.com/questions/6162789/mongodb-vs-cassandra-vs-mysql-for-real-time-advertising-platform>   Vs. MongoDB:   * <https://stackoverflow.com/questions/2892729/mongodb-vs-cassandra> |
| Costs | <https://en.wikipedia.org/wiki/Apache_Cassandra>: Free, open source |