

93 Cassandra interview questions to hire top developers

Questions

1. What is Cassandra, in simple words?
2. Can you explain why Cassandra is called a NoSQL database?
3. What are the key features of Cassandra that make it different from traditional databases?
4. Imagine you have lots of data. Why might Cassandra be a good choice to store it?
5. What does it mean when we say Cassandra is 'decentralized'?
6. What are the basic components of Cassandra's architecture?
7. Can you explain what a 'key' is in Cassandra?
8. What are column families in Cassandra? How are they structured?
9. What is data replication in Cassandra and why is it important?
10. How does Cassandra handle failures? What happens if a node goes down?
11. What is a 'consistency level' in Cassandra? Give an example.
12. Explain the difference between eventual consistency and strong consistency.
13. What is the purpose of the Cassandra Query Language (CQL)?
14. Can you give an example of a simple CQL query to insert data?
15. How do you retrieve data from Cassandra using CQL?
16. What is a 'partition key' and why is it important for querying data efficiently?
17. What are the different data types supported by Cassandra?
18. How can you update existing data in Cassandra?
19. How do you delete data from Cassandra using CQL?
20. Explain the concept of 'tuning' in Cassandra.
21. What are some common use cases for Cassandra?
22. What are the advantages of using Cassandra over other database systems for specific applications?
23. Can you explain how Cassandra handles tombstones and their impact on performance?
24. Describe the purpose and function of Bloom filters in Cassandra.
25. How does Cassandra ensure data consistency across multiple nodes in a cluster, considering different consistency levels?
26. Explain how Cassandra uses hinted handoff to maintain data durability.
27. What are the different types of counters available in Cassandra, and when would you use each?
28. Explain the process of adding or removing a node from a Cassandra cluster.
29. What are the implications of using different partitioners in Cassandra, and how do you choose the right one?
30. How do you monitor and troubleshoot Cassandra cluster performance, and what metrics are important?
31. Describe the different caching mechanisms available in Cassandra and how they improve read performance.
32. Explain how you would back up and restore a Cassandra cluster.
33. What are User Defined Functions (UDFs) in Cassandra, and how can they be useful?
34. How does Cassandra handle data compaction, and why is it necessary?
35. Explain the concept of Lightweight Transactions (LWT) in Cassandra and their use cases.
36. What are the best practices for data modeling in Cassandra to optimize performance?
37. How can you optimize Cassandra queries to improve read performance?
38. Explain the role of the commit log in Cassandra's write process.
39. What are the different consistency levels in Cassandra, and when would you choose each?
40. Describe the Cassandra Query Language (CQL) and its common commands.
41. Explain the difference between a materialized view and a secondary index in Cassandra and when to use each one.
42. How does Cassandra handle failure scenarios, such as node failures or network partitions?
43. What are the advantages and disadvantages of using Cassandra compared to other NoSQL databases?
44. Discuss strategies for handling large-scale data migrations in Cassandra without impacting application performance.
45. How does Cassandra handle consistency when data is written to multiple nodes, and what trade-offs are involved?
46. Explain the concept of hinted handoff in Cassandra and how it ensures data durability.
47. Describe the role of bloom filters in Cassandra's read path and how they improve performance.
48. What are the different types of counters in Cassandra, and when would you use each type?
49. How does Cassandra achieve high availability, and what happens when a node fails?
50. Explain the purpose of the Cassandra system keyspace and the information it contains.
51. Describe how you would monitor a Cassandra cluster to ensure optimal performance and identify potential issues.
52. What are the different types of compaction strategies available in Cassandra, and how do they impact performance?
53. How does Cassandra handle data deletion (tombstones) and what are the potential implications?
54. Explain how to configure and use virtual nodes (vnodes) in a Cassandra cluster.
55. Describe how you would back up and restore a Cassandra cluster.
56. How does Cassandra support data locality and why is it important?
57. Explain how to use the Cassandra Query Language (CQL) to perform complex queries.
58. What are the security considerations when deploying a Cassandra cluster?
59. How does Cassandra handle schema evolution, and what are the best practices for making schema changes?
60. Describe how you would troubleshoot slow queries in Cassandra.
61. How can you optimize Cassandra for write-heavy workloads?
62. Explain the concept of Lightweight Transactions (LWT) in Cassandra and their limitations.
63. How does Cassandra use SSTables and what is their role in the data lifecycle?
64. Describe how you would design a data model in Cassandra for a specific use case, considering factors like query patterns and data relationships.
65. Explain the purpose and usage of User-Defined Functions (UDFs) and User-Defined Aggregates (UDAs) in Cassandra.
66. How does Cassandra integrate with other big data technologies like Spark or Hadoop?
67. How does Cassandra handle more data than one node can hold?
68. Explain how Cassandra ensures data consistency across multiple data centers.
69. Describe the tradeoffs between eventual consistency and strong consistency in Cassandra.
70. How does Cassandra's gossip protocol contribute to cluster management and fault detection?
71. Explain the purpose and impact of tombstone in Cassandra.
72. How does Cassandra handle data replication and what are the different replication strategies available?
73. Discuss the role of the commit log and memtable in Cassandra's write path.
74. What are the implications of using different compaction strategies in Cassandra?
75. Explain how you would design a Cassandra data model for a time-series application.
76. How can you optimize Cassandra queries for performance, considering factors like filtering and indexing?
77. Describe how you would troubleshoot performance issues in a Cassandra cluster.
78. How does Cassandra support data durability and fault tolerance?
79. Explain the purpose and usage of Lightweight Transactions (LWT) in Cassandra.
80. How would you implement and manage backups and restores in a Cassandra cluster?
81. Discuss how you can monitor a Cassandra cluster and what metrics are important to track.
82. How does Cassandra's architecture differ from traditional relational databases, and what are the advantages of this difference?
83. Explain how you would handle data migrations in a Cassandra cluster without downtime.
84. Describe how Cassandra handles data consistency in the event of network partitions.
85. What are the best practices for securing a Cassandra cluster?
86. How can you integrate Cassandra with other big data technologies like Spark or Hadoop?
87. Explain how you would scale a Cassandra cluster to handle increasing data volumes and traffic.
88. Describe the steps involved in adding or removing nodes from a Cassandra cluster.
89. How does Cassandra handle concurrent writes to the same data?
90. Explain the impact of different data types on Cassandra performance.