78 Elasticsearch Interview Questions to Ask Candidates

Questions

1. What is Elasticsearch and what are its primary use cases?

2. How does Elasticsearch achieve near real-time search capabilities?

3. Can you explain what an Elasticsearch index is?

4. What are the advantages of using Elasticsearch over traditional relational databases?

5. What is a shard in Elasticsearch and why is it important?

6. How can you ensure high availability in an Elasticsearch cluster?

7. What is the role of an Elasticsearch node, and how do different types of nodes function?

8. How do you handle security in Elasticsearch?

9. What are the basic components of an Elasticsearch cluster?

10. How do you add a document to an Elasticsearch index?

11. What is the purpose of an analyzer in Elasticsearch?

12. Can you explain the difference between term-level queries and full-text queries in Elasticsearch?

13. How do you update a document in Elasticsearch?

14. What is the purpose of the 'match' query in Elasticsearch?

15. How does Elasticsearch handle data replication?

16. What is an Elasticsearch mapping?

17. How do you delete an index in Elasticsearch?

18. What is an Elasticsearch alias and how is it used?

19. Can you explain the concept of 'tokenization' in Elasticsearch?

20. How do you configure Elasticsearch to optimize search performance?

21. What is the significance of the 'match_phrase' query in Elasticsearch?

22. How would you handle large volumes of data in Elasticsearch?

23. Can you describe how Elasticsearch scoring works?

24. What is the purpose of the 'multi_match' query in Elasticsearch?

25. How do you perform a bulk operation in Elasticsearch?

26. What are 'aggregations' in Elasticsearch, and how do they work?

27. How do you monitor the health of an Elasticsearch cluster?

28. What steps would you take to troubleshoot performance issues in Elasticsearch?

29. Can you explain the purpose and function of the 'refresh interval' in Elasticsearch?

30. How would you manage a large Elasticsearch index to ensure optimal performance?

31. What are some common reasons for slow search performance in Elasticsearch and how would you troubleshoot them?

32. How do you handle data reindexing in Elasticsearch when there are schema changes?

33. Can you explain the role of 'replica shards' in Elasticsearch and how they contribute to cluster reliability?

34. How would you approach scaling an Elasticsearch cluster to handle increased load?

35. What are some best practices for designing an efficient Elasticsearch schema?

36. Can you discuss the importance of index templates in Elasticsearch and how you use them?

37. How would you implement a backup and restore strategy for an Elasticsearch cluster?

38. What methods can you use to monitor Elasticsearch performance and health?

39. Can you explain the process of creating a custom analyzer in Elasticsearch?

40. How do you optimize index settings for high-throughput scenarios?

41. Describe the differences between horizontal and vertical scaling in Elasticsearch.

42. How do you manage index lifecycle policies in Elasticsearch?

43. What strategies do you use for tuning search relevancy?

44. How would you implement a multi-tenant architecture using Elasticsearch?

45. Can you discuss how you would secure an Elasticsearch cluster in a production environment?

46. What is your approach to handling parent-child relationships in Elasticsearch?

47. How do you design efficient query patterns to minimize search latency?

48. What considerations do you take into account when setting up cross-cluster search?

49. Can you explain the importance of Elasticsearch's circuit breaker mechanism?

50. Describe a situation where you had to troubleshoot an Elasticsearch cluster issue.

51. How do you handle version conflicts during document updates?

52. What is your approach for optimizing bulk indexing operations?

53. How do you manage and monitor Elasticsearch logs effectively?

54. Can you explain how you would implement machine learning capabilities in Elasticsearch?

55. What steps do you take to ensure data consistency across a distributed Elasticsearch cluster?

56. How do you integrate Elasticsearch with other data processing systems like Apache Kafka?

57. How would you design an indexing strategy for a large dataset in Elasticsearch?

58. What are some indexing best practices you follow in Elasticsearch?

59. How do you handle different types of data in Elasticsearch indices?

60. What is your approach to managing index lifecycle policies in Elasticsearch?

61. How do you optimize indexing performance in Elasticsearch?

62. How would you handle schema changes and reindexing in Elasticsearch?

63. How do you ensure that your Elasticsearch indices are optimized for search performance?

64. What steps do you take to troubleshoot indexing issues in Elasticsearch?

65. How do you manage and optimize Elasticsearch mappings for different use cases?

66. How would you approach optimizing a slow-performing Elasticsearch query?

67. Can you explain the difference between query and filter context in Elasticsearch, and when would you use each?

68. How do you handle and optimize aggregations in Elasticsearch, especially for large datasets?

69. What strategies would you employ to optimize Elasticsearch for high-cardinality fields?

70. How would you implement and optimize a multi-field search in Elasticsearch?

71. How would you approach migrating a large-scale production Elasticsearch cluster to a newer version with minimal downtime?

72. Describe a scenario where you had to optimize Elasticsearch query performance for a high-traffic application. What steps did you take?

73. How would you design an Elasticsearch-based system to handle real-time log analysis for a large distributed application?

74. You notice that search relevance has degraded for certain queries in your e-commerce application. How would you investigate and improve the situation?

75. How would you implement a multi-tenant architecture in Elasticsearch for a SaaS application?

76. Describe how you would implement a near real-time product recommendation system using Elasticsearch.