

## Real-time, scenario-based interview questions for Apache Cassandra DB Administration, Performance Tuning & Optimization

### ◆ Apache Cassandra DBA – 500 Real-time Scenario-Based Interview Questions

#### A. Installation, Configuration & Cluster Setup (1–50)

1. How do you plan hardware sizing for a new Cassandra cluster handling high write throughput?
2. What configurations do you change first after a fresh Cassandra installation?
3. How do you set up a multi-datacenter Cassandra cluster?
4. How do you handle seed node planning in Cassandra?
5. How do you recover if all seed nodes fail at the same time?
6. What's your approach to configuring JVM heap sizes in Cassandra?
7. How do you decide num\_tokens during cluster setup?
8. How do you configure Cassandra for SSD-based storage?
9. How do you isolate commit logs, data, and caches on disks for performance?
10. How do you configure Gossip for secure inter-node communication?
11. How do you handle rack and datacenter awareness in Cassandra?
12. What configurations ensure optimal read/write consistency in production?
13. How do you secure Cassandra configuration files from unauthorized access?
14. How do you plan firewall rules for a secure Cassandra cluster?
15. How do you deploy Cassandra on Kubernetes for stateful workloads?
16. How do you configure Cassandra in a hybrid cloud setup?
17. How do you plan IP addressing for a Cassandra cluster across availability zones?
18. How do you choose between SimpleStrategy vs. NetworkTopologyStrategy for replication?
19. What happens if you set the wrong snitch configuration?
20. How do you configure Cassandra for IPv6-only environments?
21. How do you test Cassandra configuration before going live?
22. How do you ensure Cassandra starts automatically after OS reboot?
23. How do you configure Cassandra nodes for TLS encryption in transit?
24. How do you enable authentication during cluster setup?
25. How do you harden Cassandra nodes against OS-level attacks?
26. How do you configure Cassandra for mixed workloads (read-heavy + write-heavy)?
27. How do you plan schema and data model before first deployment?
28. How do you configure cassandra.yaml for high-throughput clusters?
29. How do you tune concurrent reads/writes during setup?
30. How do you decide partitioner strategy during installation?
31. How do you configure Cassandra for low-latency workloads?
32. How do you configure gossiping property file snitch vs EC2 snitch?
33. How do you avoid single point of failure in seed node selection?
34. How do you validate cluster setup after installation?
35. How do you configure key cache and row cache?
36. How do you test network connectivity across Cassandra nodes?
37. How do you configure JVM GC tuning parameters for Cassandra?

38. How do you deploy Cassandra with automation tools (Ansible/Terraform)?
39. How do you enforce configuration drift control in Cassandra clusters?
40. How do you configure Cassandra in a containerized environment?
41. How do you manage Cassandra configs with GitOps?
42. How do you configure repair service at cluster setup?
43. How do you set initial replication factors for different keyspaces?
44. How do you prepare Cassandra for disaster recovery at installation time?
45. How do you configure multi-tenant Cassandra clusters securely?
46. How do you handle Cassandra installation on mixed hardware?
47. How do you perform dry-run cluster deployments?
48. How do you configure JMX access securely in Cassandra?
49. How do you configure systemd service for Cassandra?
50. How do you verify Cassandra startup logs for successful initialization?

## **B. Administration & Daily Operations (51–150)**

51. How do you check cluster health in real-time?
52. How do you add a new node to a Cassandra cluster?
53. How do you decommission a node safely?
54. How do you handle replacing a dead node?
55. How do you change replication factor on a live cluster?
56. What's your approach to expanding a cluster in production?
57. How do you monitor Cassandra logs for operational issues?
58. How do you detect and resolve tombstone buildup?
59. How do you rotate Cassandra logs automatically?
60. How do you check consistency across replicas?
61. How do you handle unresponsive nodes in gossip ring?
62. How do you safely restart a Cassandra node in production?
63. How do you upgrade Cassandra without downtime?
64. How do you handle a mixed-version cluster during rolling upgrades?
65. How do you perform node cleanup after replication changes?
66. How do you run nodetool repair efficiently?
67. How do you automate routine Cassandra maintenance tasks?
68. How do you run nodetool decommission safely?
69. How do you detect cluster split-brain scenarios?
70. How do you configure alerting for failed repairs?
71. How do you identify dropped mutations in Cassandra logs?
72. How do you recover from corrupted SSTables?
73. How do you use nodetool scrub in production scenarios?
74. How do you archive commit logs for auditing?
75. How do you check pending compactions in Cassandra?
76. How do you check pending hints for hinted handoff?
77. How do you force a major compaction?
78. How do you automate cleanup of old snapshots?

79. How do you detect disk space pressure on nodes?
80. How do you detect high read/write latencies?
81. How do you use nodetool to check compaction history?
82. How do you manage large-scale schema migrations?
83. How do you validate schema consistency across nodes?
84. How do you check for schema disagreements?
85. How do you enable tracing for specific queries?
86. How do you disable tracing safely in production?
87. How do you configure audit logging in Cassandra?
88. How do you analyze GC logs for Cassandra?
89. How do you validate Cassandra metrics with Prometheus?
90. How do you implement role-based access control in Cassandra?
91. How do you rotate Cassandra user passwords?
92. How do you disable anonymous authentication in Cassandra?
93. How do you monitor repair lag across nodes?
94. How do you schedule repairs using Cassandra Reaper?
95. How do you manage timeouts in Cassandra configuration?
96. How do you troubleshoot coordinator node bottlenecks?
97. How do you detect hotspots in partitions?
98. How do you reshard large partitions?
99. How do you detect write amplification in SSTables?
100. How do you check for hinted handoff backlog?
101. How do you check for batch log backlog?
102. How do you monitor dropped messages in messaging service?
103. How do you check mutation failures?
104. How do you monitor gossip heartbeat signals?
105. How do you troubleshoot read repair mismatches?
106. How do you handle consistency level failures?
107. How do you monitor network latencies in Cassandra?
108. How do you validate client driver connections?
109. How do you tune thread pools for client connections?
110. How do you check request timeouts?
111. How do you adjust memtable flush thresholds?
112. How do you monitor memtable memory usage?
113. How do you detect excessive hinted handoffs?
114. How do you configure commitlog compression?
115. How do you configure commitlog archival?
116. How do you perform commitlog replay after crash?
117. How do you troubleshoot slow compaction?
118. How do you debug background thread saturation?
119. How do you troubleshoot blocked mutation stage?
120. How do you investigate coordinator timeouts?
121. How do you recover data from accidentally dropped tables?

122. How do you check cluster gossip state?
123. How do you check repair history for a keyspace?
124. How do you troubleshoot stuck repair jobs?
125. How do you resolve bootstrap failures?
126. How do you detect data imbalance across nodes?
127. How do you rebalance tokens in a cluster?
128. How do you prevent hotspotting in token allocation?
129. How do you use vnode vs. non-vnode trade-offs?
130. How do you detect duplicate data due to hints replay?
131. How do you investigate disk I/O bottlenecks?
132. How do you tune disk schedulers for Cassandra?
133. How do you troubleshoot failed compaction tasks?
134. How do you manage anti-entropy repair load?
135. How do you automate cluster-wide rolling restarts?
136. How do you check driver-side load balancing policies?
137. How do you analyze request traces for latency?
138. How do you handle corrupted commit logs?
139. How do you verify consistency level ALL reads?
140. How do you validate quorum writes across regions?
141. How do you manage schema migrations across teams?
142. How do you roll back schema changes?
143. How do you manage UDT (user-defined types) changes?
144. How do you troubleshoot schema migration failures?
145. How do you check Cassandra system keyspaces?
146. How do you monitor system\_auth keyspace health?
147. How do you troubleshoot gossip convergence delays?
148. How do you check hinted handoff replay stats?
149. How do you troubleshoot high mutation latency?
150. How do you measure coordinator load distribution?

### C. Backup, Restore & Recovery (151–200)

151. How do you perform a full cluster backup in Cassandra?
152. How do you configure incremental backups?
153. How do you restore specific keyspaces from snapshot?
154. How do you restore a single table from snapshot?
155. How do you validate integrity of Cassandra backups?
156. How do you configure backup encryption?
157. How do you transfer backups across data centers?
158. How do you automate snapshot management?
159. How do you perform PITR (point-in-time recovery)?
160. How do you recover from accidental TRUNCATE?
161. How do you restore data from commit logs?
162. How do you validate restore performance?

163. How do you test DR restores regularly?
164. How do you back up schema only without data?
165. How do you configure backup retention policies?
166. How do you optimize backup size?
167. How do you throttle backup bandwidth?
168. How do you back up large partitions efficiently?
169. How do you back up multi-region clusters?
170. How do you integrate Cassandra backup with cloud storage (S3/GCS)?
171. How do you recover from disk corruption using backups?
172. How do you perform backup under heavy cluster load?
173. How do you recover if backup metadata is corrupted?
174. How do you automate restore tests?
175. How do you restore backups into test environments?
176. How do you migrate clusters using backups?
177. How do you back up system keyspaces?
178. How do you detect backup failures?
179. How do you configure alerting for failed backups?
180. How do you manage encryption keys for backup data?
181. How do you validate backup consistency across replicas?
182. How do you restore data with partial replica loss?
183. How do you restore with a different replication factor?
184. How do you recover when cluster topology changes post-backup?
185. How do you validate checksums of backup files?
186. How do you recover if incremental backups are missing?
187. How do you manage large backup catalogs?
188. How do you reduce restore times for large datasets?
189. How do you integrate backup tools (Medusa, etc.)?
190. How do you restore specific rows from backup?
191. How do you encrypt commit logs for secure backups?
192. How do you automate PITR with binlog-like systems?
193. How do you throttle restore process?
194. How do you back up Cassandra running in Docker?
195. How do you test restores for SLA compliance?
196. How do you migrate Cassandra to cloud using backups?
197. How do you monitor backup storage utilization?
198. How do you recover when commit log + snapshot mismatch?
199. How do you configure multi-datacenter consistent backups?
200. How do you validate restore didn't cause data loss?

#### **D. Performance Tuning & Optimization (201–350)**

201. How do you identify slow queries in Cassandra?
202. How do you optimize read-heavy workloads?
203. How do you optimize write-heavy workloads?

204. How do you tune compaction strategies?
205. How do you optimize LeveledCompactionStrategy?
206. How do you optimize SizeTieredCompactionStrategy?
207. How do you select TimeWindowCompactionStrategy?
208. How do you detect compaction backlog issues?
209. How do you tune memtable size for workload optimization?
210. How do you tune commitlog size for throughput?
211. How do you detect and fix GC pauses?
212. How do you tune heap size for performance?
213. How do you reduce tombstone impact on queries?
214. How do you tune read\_request\_timeout\_in\_ms?
215. How do you tune write\_request\_timeout\_in\_ms?
216. How do you optimize batch writes in Cassandra?
217. How do you reduce latency in quorum reads?
218. How do you detect hotspot partitions?
219. How do you optimize data modeling for wide rows?
220. How do you detect unbalanced partition sizes?
221. How do you use materialized views effectively?
222. How do you optimize secondary indexes?
223. How do you detect query timeouts?
224. How do you configure caching for reads?
225. How do you tune row cache usage?
226. How do you monitor page cache effectiveness?
227. How do you detect coordinator overload?
228. How do you optimize coordinator selection?
229. How do you reduce cross-DC query latency?
230. How do you optimize hinted handoff settings?
231. How do you tune repair frequency?
232. How do you tune streaming during bootstrap?
233. How do you reduce repair resource usage?
234. How do you throttle streaming during bootstrap?
235. How do you reduce read repair latency?
236. How do you tune speculative retries?
237. How do you tune coordinator timeouts?
238. How do you optimize prepared statements?
239. How do you optimize pagination queries?
240. How do you tune consistency levels for latency vs. availability?
241. How do you use LOCAL\_QUORUM effectively?
242. How do you configure token-aware load balancing?
243. How do you optimize driver connection pools?
244. How do you detect driver-side bottlenecks?
245. How do you configure driver request throttling?
246. How do you analyze query profiles for performance?

247. How do you debug queries with tracing?
248. How do you monitor coordinator vs. replica load?
249. How do you optimize user-defined functions?
250. How do you detect slow user-defined aggregates?
251. How do you optimize GC-friendly data modeling?
252. How do you avoid clustering column hotspots?
253. How do you manage large partition keys?
254. How do you balance partition key cardinality?
255. How do you optimize schema for time-series data?
256. How do you manage TTL impact on performance?
257. How do you detect TTL expiry storms?
258. How do you throttle TTL deletions?
259. How do you optimize compaction for TTL-heavy tables?
260. How do you optimize SSTable size for performance?
261. How do you monitor bloom filter false positives?
262. How do you tune bloom filters for reads?
263. How do you tune key cache for reads?
264. How do you detect cache thrashing?
265. How do you preheat caches after restart?
266. How do you tune concurrency for compaction threads?
267. How do you configure parallelism in repairs?
268. How do you reduce lock contention in Cassandra threads?
269. How do you detect high heap usage in caches?
270. How do you optimize thread pool sizes?
271. How do you configure async thread pools?
272. How do you tune mutation stage threads?
273. How do you optimize read stage threads?
274. How do you optimize memtable flush writers?
275. How do you optimize compaction executor threads?
276. How do you detect thread pool saturation?
277. How do you avoid large batch mutations?
278. How do you tune speculative retry strategies?
279. How do you optimize cross-shard queries?
280. How do you tune SSTable compression?
281. How do you detect compression ratio inefficiencies?
282. How do you optimize compression chunk size?
283. How do you analyze read amplification?
284. How do you optimize write amplification?
285. How do you detect compaction IO bottlenecks?
286. How do you tune disk IO schedulers?
287. How do you detect coordinator-to-replica skew?
288. How do you optimize data distribution across nodes?
289. How do you tune vnode count for performance?



290. How do you validate token distribution fairness?
291. How do you monitor gossip convergence delays?
292. How do you detect high dropped messages?
293. How do you tune internode compression?
294. How do you optimize internode messaging threads?
295. How do you configure internode TCP settings?
296. How do you detect high internode latencies?
297. How do you optimize batchlog replay?
298. How do you detect batchlog backlog?
299. How do you throttle batchlog replay?
300. How do you monitor hinted handoff delays?
301. How do you tune hinted handoff performance?
302. How do you monitor mutation backpressure?
303. How do you tune backpressure thresholds?
304. How do you optimize anti-entropy repairs?
305. How do you monitor repair streaming throughput?
306. How do you tune repair parallelism?
307. How do you configure incremental repairs?
308. How do you detect repair-induced latency?
309. How do you tune stream\_throughput\_outbound\_megabits\_per\_sec?
310. How do you reduce cross-DC repair costs?
311. How do you detect repair conflicts?
312. How do you optimize schema for analytics queries?
313. How do you tune read-before-write scenarios?
314. How do you optimize counters in Cassandra?
315. How do you detect counter inconsistencies?
316. How do you reduce counter replication overhead?
317. How do you optimize time-to-live workloads?
318. How do you optimize multi-tenant workloads?
319. How do you throttle tenant workloads fairly?
320. How do you optimize Cassandra for IoT use cases?
321. How do you handle bursty workloads in Cassandra?
322. How do you configure rate limiting in Cassandra clients?
323. How do you detect GC stalls under high load?
324. How do you tune garbage collectors for low latency?
325. How do you use G1GC effectively in Cassandra?
326. How do you optimize heap for mixed workloads?
327. How do you monitor JVM heap fragmentation?
328. How do you tune young gen size?
329. How do you tune old gen size?
330. How do you detect frequent full GCs?
331. How do you monitor Cassandra thread dump for bottlenecks?
332. How do you analyze flame graphs for Cassandra performance?



- 333. How do you optimize read path for locality?
- 334. How do you optimize write path for low latency?
- 335. How do you tune coordinator thread scheduling?
- 336. How do you optimize request batching strategies?
- 337. How do you detect read-before-write penalties?
- 338. How do you optimize retry policies?
- 339. How do you tune driver-side speculative execution?
- 340. How do you monitor prepared statement cache hit ratio?
- 341. How do you tune prepared statement cache size?
- 342. How do you analyze statement idempotency for retries?
- 343. How do you optimize user-defined functions performance?
- 344. How do you optimize multi-region workload routing?
- 345. How do you configure consistency levels for analytics workloads?
- 346. How do you optimize Cassandra with Spark connector?
- 347. How do you troubleshoot Spark-Cassandra connector bottlenecks?
- 348. How do you optimize Kafka-to-Cassandra ingestion?
- 349. How do you handle backpressure in Kafka → Cassandra pipelines?
- 350. How do you tune CDC (change data capture) performance?

#### **E. Security, Auditing & Compliance (351–400)**

- 351. How do you enforce TLS for all Cassandra clients?
- 352. How do you rotate TLS certificates in Cassandra?
- 353. How do you implement client-to-node encryption?
- 354. How do you secure internode communication with TLS?
- 355. How do you rotate internode TLS certificates?
- 356. How do you enforce role-based access for multi-tenant clusters?
- 357. How do you enable audit logging in Cassandra?
- 358. How do you capture failed login attempts?
- 359. How do you audit schema changes?
- 360. How do you enforce password complexity for Cassandra users?
- 361. How do you configure LDAP authentication for Cassandra?
- 362. How do you configure Kerberos authentication?
- 363. How do you rotate admin credentials securely?
- 364. How do you monitor access logs for anomalies?
- 365. How do you enforce encryption at rest for SSTables?
- 366. How do you encrypt commit logs at rest?
- 367. How do you enable disk-level encryption?
- 368. How do you implement key management service for Cassandra?
- 369. How do you rotate encryption keys for SSTables?
- 370. How do you audit node-to-node SSL handshake failures?
- 371. How do you detect privilege escalation attempts?
- 372. How do you enforce least privilege access in Cassandra?
- 373. How do you audit Cassandra schema migrations?

- 374. How do you monitor client IPs connecting to Cassandra?
- 375. How do you block malicious IPs?
- 376. How do you configure firewall for Cassandra security?
- 377. How do you integrate Cassandra with SIEM tools?
- 378. How do you monitor suspicious schema modifications?
- 379. How do you prevent unauthorized JMX access?
- 380. How do you configure SSL for JMX?
- 381. How do you audit JMX operations?
- 382. How do you enable audit trail for login/logout events?
- 383. How do you implement secure auditing for GDPR compliance?
- 384. How do you handle Cassandra security patches?
- 385. How do you check cluster for CVE vulnerabilities?
- 386. How do you manage Cassandra upgrades with security fixes?
- 387. How do you enforce compliance with PCI DSS?
- 388. How do you enforce HIPAA compliance in Cassandra?
- 389. How do you mask sensitive data in Cassandra?
- 390. How do you encrypt sensitive columns in Cassandra?
- 391. How do you monitor access to sensitive keyspaces?
- 392. How do you restrict schema changes to admin users only?
- 393. How do you configure data redaction in Cassandra logs?
- 394. How do you detect brute-force login attempts?
- 395. How do you block repeated failed login attempts?
- 396. How do you integrate Cassandra audit logs with ELK stack?
- 397. How do you configure Cassandra firewall rules in AWS?
- 398. How do you monitor Cassandra TLS handshake latencies?
- 399. How do you validate certificate expiry in Cassandra?
- 400. How do you audit multi-datacenter connections?

#### **F. High Availability, DR & Scaling (401–500)**

- 401. How do you configure Cassandra for cross-region replication?
- 402. How do you plan replication factor for HA workloads?
- 403. How do you configure LOCAL\_QUORUM for multi-DC HA?
- 404. How do you test failover across datacenters?
- 405. How do you simulate node failure in production?
- 406. How do you recover from full DC outage?
- 407. How do you configure Cassandra for active-active clusters?
- 408. How do you plan RPO and RTO for Cassandra clusters?
- 409. How do you validate DR strategy for Cassandra?
- 410. How do you failover clients across DCs?
- 411. How do you monitor cross-DC replication lag?
- 412. How do you optimize inter-DC replication traffic?
- 413. How do you configure hinted handoff across DCs?
- 414. How do you handle write failures in multi-DC clusters?

415. How do you simulate region-wide Cassandra failures?
416. How do you design Cassandra clusters for hybrid cloud?
417. How do you configure Cassandra clusters across AWS regions?
418. How do you configure Cassandra clusters across Azure regions?
419. How do you configure Cassandra clusters across GCP regions?
420. How do you scale Cassandra horizontally?
421. How do you test cluster scaling in pre-prod?
422. How do you validate performance after node addition?
423. How do you troubleshoot rebalance failures after scaling?
424. How do you automate scaling with Kubernetes?
425. How do you tune consistency levels during scaling?
426. How do you plan rack-aware scaling?
427. How do you handle heterogeneous hardware scaling?
428. How do you migrate from on-prem to cloud Cassandra?
429. How do you test zero-downtime migrations?
430. How do you perform live cluster cloning?
431. How do you handle multi-tenancy during scaling?
432. How do you simulate rolling network failures?
433. How do you simulate disk failures in Cassandra nodes?
434. How do you simulate split-brain scenarios?
435. How do you validate HA after node replacements?
436. How do you manage leaderless HA in Cassandra?
437. How do you configure driver-side HA for apps?
438. How do you tune request consistency for HA apps?
439. How do you configure retry policies for HA workloads?
440. How do you recover from mass node failure?
441. How do you restore quorum after multiple node failures?
442. How do you configure repair strategy for HA workloads?
443. How do you validate consistency after DR restores?
444. How do you validate HA after schema changes?
445. How do you plan load balancing for multi-tenant HA?
446. How do you test HA in rolling upgrade scenarios?
447. How do you design Cassandra HA for IoT workloads?
448. How do you test HA under flash crowd scenarios?
449. How do you recover from accidental data truncation in HA cluster?
450. How do you optimize HA workloads for quorum reads/writes?
451. How do you validate HA during massive write bursts?
452. How do you configure async replication for HA?
453. How do you ensure data consistency in eventual HA models?
454. How do you simulate Cassandra node gossip failures?
455. How do you monitor cross-DC repair overhead?
456. How do you configure HA with heterogeneous consistency levels?
457. How do you test app failover under Cassandra failover?

458. How do you monitor HA metrics in Cassandra clusters?
459. How do you validate HA for Spark-Cassandra connector?
460. How do you validate HA for Kafka-to-Cassandra pipelines?
461. How do you design HA for change data capture (CDC)?
462. How do you configure multi-region read replicas?
463. How do you design Cassandra for read-intensive HA?
464. How do you design Cassandra for write-intensive HA?
465. How do you scale Cassandra CDC workloads?
466. How do you validate CDC replication consistency?
467. How do you test CDC during node failures?
468. How do you configure tiered storage in HA clusters?
469. How do you migrate tiered storage HA setups?
470. How do you configure HA with Kubernetes operators?
471. How do you manage HA Cassandra in OpenShift?
472. How do you design active-passive Cassandra setups?
473. How do you design active-active Cassandra setups?
474. How do you simulate multi-DC failover in pre-prod?
475. How do you validate RTO under failover conditions?
476. How do you validate RPO under failover conditions?
477. How do you test partial replica availability?
478. How do you design HA for financial workloads?
479. How do you validate HA for real-time analytics?
480. How do you validate HA for session management systems?
481. How do you simulate data center isolation?
482. How do you recover from cluster-wide gossip failure?
483. How do you validate HA for batch analytics pipelines?
484. How do you simulate WAN link failures?
485. How do you configure HA for hybrid multi-cloud?
486. How do you simulate rolling DC outages?
487. How do you configure HA monitoring dashboards?
488. How do you implement SLA tracking for HA workloads?
489. How do you configure alerts for HA SLA breaches?
490. How do you validate DR readiness quarterly?
491. How do you automate HA failover testing?
492. How do you simulate cloud provider outages?
493. How do you configure Cassandra in chaos engineering tests?
494. How do you monitor cluster stability during chaos tests?
495. How do you simulate compaction storms under HA?
496. How do you test repair workloads under failover?
497. How do you validate quorum consistency under load?
498. How do you validate latency SLAs under HA workloads?
499. How do you design Cassandra HA for healthcare workloads?
500. How do you document HA/DR playbooks for Cassandra teams?

## Another Set of Questions:

Perfect 🙌 Since you asked for **500 realtime scenario-based interview questions on Apache Cassandra DB Administration, Performance Tuning & Optimisation** — all in one go, no chunks, no docs, and without stopping — here's the **full set from 1 to 500**:

---

### Apache Cassandra DB — 500 Realtime Scenario Based Interview Questions

#### Cluster Setup & Administration (1–100)

1. Your Cassandra cluster has 10 nodes across 3 data centers. How would you ensure replication is balanced?
2. A node frequently goes down during peak hours. How do you identify whether it's hardware, JVM, or Cassandra-related?
3. How do you plan data center awareness while configuring a new keyspace?
4. A client demands 99.99% uptime. What Cassandra configurations would you focus on?
5. How do you safely decommission a node from a production Cassandra cluster?
6. During a rolling upgrade, one node refuses to join the ring. How do you troubleshoot?
7. What's your approach if nodetool status shows one node as "UN" but it's not serving queries?
8. How do you configure gossip protocols for secure multi-data center communication?
9. Explain the process of bootstrapping a new Cassandra node into an existing ring.
10. A disk has filled up on one node. How do you free up space without affecting cluster availability?
11. How would you use "auto\_bootstrap" in Cassandra?
12. How do you handle consistency levels across multiple regions for critical financial transactions?
13. When would you recommend using "NetworkTopologyStrategy" instead of "SimpleStrategy"?
14. Explain how you'd enable incremental repair in Cassandra.
15. What steps would you take if the cluster consistently reports "hinted handoff queue is full"?
16. A node crashes during compaction. How do you recover without data loss?
17. How do you perform routine health checks on Cassandra clusters?
18. Describe your approach for configuring seed nodes in a 50-node cluster.
19. How would you secure Cassandra nodes in a hybrid cloud setup?
20. When should you perform a rolling restart of Cassandra services?
21. What's the impact of adding a new data center to an existing cluster?
22. How do you configure authentication & authorization in Cassandra?
23. A client requests fine-grained access controls. How do you set them up?
24. Explain how to manage schema migrations across multiple environments.
25. During schema migration, how do you avoid "schema disagreement"?
26. If you need to migrate from 3.11 to 4.1, what steps would you take?
27. A node is repeatedly marked "down" by gossip. How do you investigate?
28. How do you decide the right JVM heap size for Cassandra nodes?
29. What monitoring dashboards are critical for Cassandra production environments?
30. How do you handle a corrupted commit log?
31. A node fails during major compaction. What's your recovery path?
32. How do you configure and test incremental backups in Cassandra?
33. How do you detect and resolve tombstone accumulation?
34. Explain a situation where you had to forcefully remove a dead node from the ring.

35. What's your approach to managing SSTable bloat?
36. How do you ensure repair operations don't impact SLAs?
37. How do you handle a split-brain situation in Cassandra clusters?
38. How do you scale Cassandra horizontally without disrupting workloads?
39. Which metrics do you monitor to detect GC pressure?
40. How do you configure SSL/TLS for Cassandra inter-node communication?
41. A node has an out-of-sync schema. How do you resolve it?
42. Explain steps to recover a node after disk corruption.
43. What's your approach if "nodetool repair" takes forever?
44. How do you plan capacity for a cluster storing 200TB of data?
45. How do you configure multiple keyspaces with different consistency requirements?
46. Explain how you'd back up and restore a single keyspace.
47. How do you configure DSE security in Cassandra?
48. What steps would you take if queries suddenly start failing with "unavailable exception"?
49. How do you validate your cluster after a full repair?
50. What's the role of Hinted Handoff and when should it be disabled?
51. A node consistently shows "streaming error". How do you resolve it?
52. How do you detect hotspots in your cluster?
53. What's your approach to replacing a dead seed node?
54. Explain the steps to enable audit logging.
55. How do you clean up "zombie" hints on Cassandra nodes?
56. A cluster has too many pending compactions. What actions do you take?
57. How do you protect Cassandra from disk I/O bottlenecks?
58. How do you enable encryption for data at rest?
59. What's your procedure for restoring snapshots on a live cluster?
60. How do you schedule nodetool repairs in a large environment?
61. Explain how you'd recover from a mistaken TRUNCATE query.
62. How do you monitor slow queries in Cassandra?
63. A client requests GDPR-compliant data erasure. How do you implement it?
64. How do you safely upgrade drivers during live workloads?
65. Which performance metrics do you alert on in Grafana/Prometheus?
66. How do you perform schema migration testing?
67. Explain the difference between full vs incremental repair.
68. What's the impact of misconfigured snitches?
69. How do you simulate node failures for DR drills?
70. How do you enable row-level security in Cassandra?
71. A node has 90% heap usage. What actions do you take?
72. How do you configure compaction strategies for time-series data?
73. Explain your strategy for handling cross-region replication.
74. A keyspace needs both high consistency and low latency. How do you balance this?
75. How do you validate backups regularly?
76. How do you implement auditing for sensitive columns?
77. What's the impact of over-provisioning vnodes?



78. A schema change triggers cluster-wide instability. How do you mitigate?
79. How do you handle mismatched partitioners in a cluster?
80. Which nodetool commands do you use for troubleshooting?
81. Explain the process for enabling multi-tenancy in Cassandra.
82. How do you reduce tombstone-related query timeouts?
83. A client demands encryption at both client and inter-node levels. How do you configure it?
84. How do you troubleshoot inconsistent read results?
85. How do you perform an emergency patch upgrade?
86. A keyspace must be replicated to only 2 out of 3 data centers. How do you configure?
87. How do you optimize commit log archiving?
88. What steps do you take for OS tuning before deploying Cassandra?
89. How do you configure authentication with LDAP/AD?
90. What's the role of gc\_grace\_seconds?
91. How do you troubleshoot high pending hints?
92. How do you plan compaction during peak workloads?
93. A client complains of high latency during repairs. How do you address it?
94. How do you configure dual-write auditing?
95. A snapshot restore corrupts SSTables. How do you handle it?
96. Explain how to secure Cassandra with Kerberos.
97. How do you prepare for a 10x workload increase?
98. What's your strategy for distributing workloads evenly across nodes?
99. How do you tune OS-level file descriptors for Cassandra?
100. A node keeps rejoining with a new host ID. How do you fix it?

Great 👍 Thanks for confirming.

Here is the **complete continuation from 101 to 500** for **Apache Cassandra DB Administration, Performance Tuning & Optimisation (Scenario Based)** in one go, without stopping:

---

## Apache Cassandra DB — 500 Realtime Scenario Based Interview Questions (Continued)

---

### Performance Tuning & Optimization (101–200)

101. A query that used to run in 10ms now takes 2s. How do you investigate?
102. How do you tune JVM garbage collection in Cassandra?
103. A table is suffering from tombstone timeouts. How do you optimize it?
104. How do you decide between Leveled Compaction Strategy (LCS) vs Size-Tiered Compaction Strategy (STCS)?
105. What's the impact of changing compaction\_throughput\_mb\_per\_sec?
106. How do you optimize commit log sync settings for high-write workloads?
107. Explain your approach to tuning read latency in Cassandra.
108. How do you resolve hotspots caused by uneven partition distribution?
109. A batch insert job is failing. How do you tune it for large workloads?
110. How do you optimize queries suffering from coordinator overload?
111. How do you tune the number of concurrent\_reads and concurrent\_writes?
112. Which OS-level tunings are most critical for Cassandra performance?



113. How do you analyze SSTable read patterns?
114. Explain how you'd reduce GC pressure during high compaction.
115. How do you tune memtable\_flush\_writers?
116. A cluster experiences high disk I/O. How do you balance compactions?
117. What's the role of speculative retries in query performance?
118. How do you configure page size for queries to avoid timeouts?
119. A table has large partitions. How do you refactor schema to improve reads?
120. What's the impact of Bloom filter false positives on read latency?
121. How do you optimize coordinator performance in multi-DC clusters?
122. How do you configure caching for read-heavy workloads?
123. How do you detect and optimize queries using ALLOW FILTERING?
124. Explain steps to reduce read repair impact on latency.
125. A query is scanning entire partitions. How do you rewrite it?
126. How do you configure throttling for heavy workloads?
127. Which parameters do you tune for write-heavy time-series workloads?
128. Explain how materialized views affect performance.
129. How do you optimize wide-row queries?
130. What's your strategy for improving performance in IoT workloads?
131. How do you balance compaction and repair performance?
132. A cluster suffers from heap pressure. How do you tune it?
133. What's the role of key\_cache\_size\_in\_mb?
134. How do you tune row\_cache\_size\_in\_mb?
135. How do you identify slow nodes in a cluster?
136. A high-volume read query always hits disk. How do you optimize?
137. What's the role of read\_repair\_chance?
138. How do you tune consistency levels for latency-sensitive workloads?
139. A cluster shows high dropped messages. How do you handle it?
140. How do you tune gossip settings for performance?
141. Explain how adaptive bloom filters improve read efficiency.
142. A node is always behind in hinted handoffs. How do you fix it?
143. How do you tune cassandra.yaml for multi-threaded workloads?
144. What's the role of batchlog\_replay\_throttle\_in\_kb?
145. How do you tune Cassandra for hybrid read/write workloads?
146. A single partition grows beyond 1GB. How do you redesign schema?
147. Explain your approach to optimizing queries with secondary indexes.
148. How do you tune SSTable compression for performance?
149. What's your strategy for handling large data imports efficiently?
150. How do you optimize time-to-live (TTL) settings for write-heavy data?
151. How do you configure hints\_directory for performance?
152. A client app causes high read timeouts. How do you troubleshoot?
153. How do you optimize thread pools in Cassandra?
154. What's the impact of setting speculative\_retry to ALWAYS?
155. How do you detect if latency issues are due to compaction?

156. Explain the impact of bloom\_filter\_fp\_chance on read performance.
157. A cluster has high CPU usage. How do you narrow down the cause?
158. How do you optimize for mixed OLTP + analytics workloads?
159. Which parameters affect memtable flushing performance?
160. How do you detect partition key skew in workloads?
161. How do you optimize Cassandra for SSDs vs spinning disks?
162. What's the role of commitlog\_sync\_batch\_window\_in\_ms?
163. How do you minimize tombstone warnings in logs?
164. A table has hundreds of secondary indexes. How do you optimize?
165. How do you prevent hotspots from counters?
166. Which compaction settings would you tune for heavy TTL workloads?
167. How do you tune cluster for write-heavy log ingestion use cases?
168. A query returns stale data. How do you ensure freshness?
169. How do you optimize consistency levels for multi-region workloads?
170. A node's query performance degrades after restart. How do you investigate?
171. How do you tune num\_tokens for vnode distribution?
172. What's the role of dynamic\_snitch\_badness\_threshold?
173. How do you troubleshoot coordinator-side latency?
174. A query performs poorly due to large tombstones. How do you resolve?
175. How do you design partition keys for balanced read/write performance?
176. How do you detect GC issues in real time?
177. What's the impact of increasing compaction\_large\_partition\_warning\_threshold\_mb?
178. How do you tune for latency-sensitive gaming workloads?
179. How do you optimize queries on composite partition keys?
180. A repair job hogs CPU. How do you tune repair\_parallelism?
181. What's the role of memtable\_flush\_queue\_size?
182. How do you monitor bloom filter usage?
183. A high-traffic node drops connections. How do you fix it?
184. How do you optimize materialized views with high write amplification?
185. What's the impact of increasing key\_cache\_save\_period?
186. How do you optimize disk I/O scheduler for Cassandra?
187. How do you handle query timeouts due to coordinator overload?
188. Which repair strategy do you choose for large datasets?
189. How do you optimize data modeling for time-series events?
190. A client demands 5ms read latency. What tuning parameters do you adjust?
191. How do you tune retry\_policy in client drivers?
192. A table grows very large due to TTL churn. How do you redesign it?
193. What's the impact of increasing internode\_compression?
194. How do you optimize for cold reads in Cassandra?
195. Explain tuning for high-throughput batch jobs.
196. How do you prevent hotspots when using TimeUUID?
197. What's your approach to optimizing cluster under network latency?
198. How do you detect queries causing coordinator overload?

- 199. How do you tune hinted handoff for low-latency SLAs?
- 200. Which tools do you use for workload benchmarking?

### Schema Design & Query Optimization (201–300)

- 201. A table frequently causes "unavailable exception". How do you redesign schema?
- 202. How do you model schema for 1-to-many relationships in Cassandra?
- 203. How do you enforce uniqueness without global constraints?
- 204. What's the impact of large partitions on query performance?
- 205. How do you optimize schema for hierarchical data?
- 206. How do you handle schema evolution with minimal downtime?
- 207. Explain your approach for schema design in IoT workloads.
- 208. How do you model schema for multi-tenant SaaS apps?
- 209. How do you redesign schema if queries rely heavily on ALLOW FILTERING?
- 210. How do you optimize schema for time-windowed data?
- 211. A client wants ACID-like guarantees. How do you design schema?
- 212. Explain how you'd handle schema migration across DCs.
- 213. How do you detect and fix partition key skew?
- 214. How do you optimize queries on clustering columns?
- 215. A query uses multiple secondary indexes. How do you optimize schema?
- 216. How do you optimize queries that scan wide rows?
- 217. What's the impact of using composite partition keys?
- 218. How do you model schema for product catalogs with multiple filters?
- 219. A query frequently uses IN clause. How do you optimize schema?
- 220. How do you handle schema design for range queries?
- 221. Explain schema modeling for event sourcing.
- 222. A query fails due to tombstone overflows. How do you optimize schema?
- 223. How do you optimize schema for geo-spatial queries?
- 224. What's the role of denormalization in Cassandra schema design?
- 225. How do you optimize schema for OLAP queries?
- 226. How do you balance query flexibility vs performance in schema design?
- 227. A client requires dynamic attributes. How do you model schema?
- 228. How do you handle schema for high cardinality fields?
- 229. A query must filter across multiple columns. How do you design schema?
- 230. How do you reduce tombstone creation through schema redesign?
- 231. Explain schema patterns for messaging applications.
- 232. How do you design schema for inventory management?
- 233. How do you optimize schema for recommendation systems?
- 234. What's your schema strategy for user activity logs?
- 235. A query always causes full partition scans. How do you redesign schema?
- 236. How do you optimize schema for sensor data streams?
- 237. How do you implement time-bucketed schema design?
- 238. Explain schema patterns for order processing.
- 239. A query always filters on non-partition keys. How do you fix schema?

240. How do you design schema for leaderboards?
241. What's your approach to schema for fraud detection?
242. How do you handle schema for auditing systems?
243. How do you optimize schema for batch inserts?
244. Explain schema design for e-commerce browsing history.
245. A query mixes multiple access patterns. How do you redesign schema?
246. How do you optimize schema for content management?
247. What's your approach for schema in microservices?
248. A query times out due to large clustering key range. How do you redesign?
249. How do you optimize schema for chat applications?
250. How do you design schema for financial transactions?
251. Explain schema for time-series stock data.
252. A query must always return latest record. How do you optimize schema?
253. How do you avoid hot partitions in schema design?
254. How do you design schema for IoT sensor rollups?
255. What's your schema strategy for versioned data?
256. How do you optimize schema for historical reporting?
257. A query needs to sort data globally. How do you redesign schema?
258. How do you model schema for clickstream analytics?
259. What's your schema design for gaming leaderboards?
260. How do you handle schema for regulatory compliance?
261. A query mixes equality and range filters. How do you optimize?
262. How do you optimize schema for multi-attribute filtering?
263. How do you implement write-optimized schema design?
264. How do you optimize schema for distributed joins (via app side)?
265. A schema migration caused downtime. How do you redesign migration strategy?
266. How do you handle schema for hierarchical product categories?
267. Explain schema design for airline reservation systems.
268. A query must join multiple datasets. How do you redesign schema?
269. How do you optimize schema for OLTP workloads?
270. A client needs query flexibility. How do you design schema without degrading performance?
271. How do you avoid over-partitioning?
272. What's the impact of schema with wide clustering keys?
273. How do you design schema for multi-region consistency?
274. A query uses LIKE patterns. How do you optimize schema?
275. How do you optimize schema for healthcare records?
276. What's your schema design for payment processing?
277. How do you handle schema for IoT device metadata?
278. How do you optimize schema for shipping/logistics?
279. Explain schema patterns for social networking apps.
280. A query retrieves trending items. How do you model schema?
281. How do you optimize schema for real-time dashboards?
282. A client requires historical + real-time access. How do you design schema?

- 283. How do you optimize schema for HR systems?
- 284. What's your schema strategy for machine learning pipelines?
- 285. How do you design schema for telecom call records?
- 286. A query retrieves time-bucketed aggregates. How do you optimize schema?
- 287. How do you optimize schema for document metadata?
- 288. Explain schema for log aggregation systems.
- 289. How do you optimize schema for workflow systems?
- 290. How do you handle schema for personalization engines?
- 291. A query aggregates over billions of rows. How do you optimize schema?
- 292. How do you design schema for ticketing systems?
- 293. How do you optimize schema for IoT device telemetry?
- 294. How do you handle schema for compliance data retention?
- 295. What's your schema approach for catalog search?
- 296. A schema redesign requires downtime. How do you avoid it?
- 297. How do you design schema for customer profiles?
- 298. How do you optimize schema for ad tech platforms?
- 299. Explain schema design for content recommendation.
- 300. A query requires multi-table consistency. How do you design schema?

#### **Buffer/Thread Tuning & Workload Balancing (301–400)**

- 301. How do you configure thread pools for mixed workloads?
- 302. A cluster has high mutation stage latency. How do you tune?
- 303. How do you balance workloads across DCs?
- 304. What's the role of concurrent\_compactors?
- 305. How do you tune buffer sizes for network throughput?
- 306. How do you troubleshoot blocked flush writers?
- 307. How do you configure workload isolation for analytics vs OLTP?
- 308. A cluster shows dropped mutations. How do you tune?
- 309. How do you detect unbalanced vnode token assignments?
- 310. How do you optimize for NUMA-aware deployments?
- 311. How do you tune heap\_new\_size for better thread performance?
- 312. How do you configure Cassandra for bursty workloads?
- 313. A cluster experiences long GC pauses. How do you tune thread pools?
- 314. What's your strategy for managing multiple workloads on the same cluster?
- 315. How do you tune for CPU-intensive workloads?
- 316. A node consistently lags behind others. How do you balance load?
- 317. How do you tune inbound vs outbound connections?
- 318. How do you use workload simulators for balancing?
- 319. A cluster sees uneven request distribution. How do you debug?
- 320. How do you configure commitlog segment sizes?
- 321. How do you balance read-heavy vs write-heavy workloads?
- 322. How do you tune for latency-critical workloads like gaming?
- 323. What's your approach for isolating analytics queries?

324. How do you detect and fix thread starvation?
325. How do you balance workloads during node decommission?
326. A workload spikes at predictable times. How do you tune Cassandra?
327. How do you optimize compaction thread utilization?
328. Explain how you balance coordinator load across nodes.
329. A client app sends too many small writes. How do you optimize?
330. How do you configure memtable allocations for workloads?
331. What's the impact of increasing netty thread pools?
332. How do you handle cluster overload from sudden traffic spikes?
333. How do you distribute workloads in multi-region clusters?
334. How do you optimize OS scheduler for Cassandra workloads?
335. How do you detect imbalance in partition distribution?
336. A node runs hotter than others. How do you troubleshoot?
337. How do you tune message queues for stability?
338. How do you configure cassandra-env.sh for optimal threading?
339. A cluster frequently experiences GC storms. How do you tune threads?
340. How do you ensure fair load balancing between clients?
341. How do you tune kernel networking for Cassandra?
342. What's your approach for tuning disk I/O threads?
343. A node lags in streaming operations. How do you fix?
344. How do you configure OS hugepages for Cassandra?
345. How do you tune Cassandra for mixed CPU/memory workloads?
346. A cluster shows high latency under low CPU usage. How do you investigate?
347. How do you configure compaction for parallel workloads?
348. How do you optimize coordinator batch threads?
349. A cluster shows uneven distribution of write requests. How do you balance?
350. How do you configure thread pools for cross-region workloads?
351. How do you optimize backpressure handling in Cassandra?
352. A node crashes under heavy client load. How do you tune it?
353. How do you balance workloads during rolling upgrades?
354. How do you configure off-heap memory tuning?
355. What's your approach to balancing repair workload?
356. How do you configure thread pools for speculative retries?
357. A query causes blocking across nodes. How do you tune?
358. How do you balance workloads across client drivers?
359. How do you optimize Cassandra for hybrid transactional + batch jobs?
360. How do you tune memtable\_cleanup\_threshold?
361. A cluster shows frequent "unavailable" exceptions. How do you tune?
362. How do you optimize OS networking buffers?
363. How do you balance workloads in Kubernetes Cassandra deployments?
364. How do you tune netty\_max\_direct\_memory?
365. A node shows very high pending tasks. How do you reduce it?
366. How do you configure load balancing policies in drivers?



367. How do you detect and resolve CPU-bound workloads?
368. How do you configure off-heap bloom filter cache?
369. What's the role of max\_hints\_delivery\_threads?
370. How do you configure parallelism for streaming data?
371. A repair operation hogs bandwidth. How do you tune?
372. How do you configure memory allocator for Cassandra threads?
373. How do you balance workloads during resharding?
374. How do you configure Cassandra for write-heavy OLTP?
375. A node constantly fails compactions. How do you optimize threads?
376. How do you configure read\_repair\_threads?
377. A workload saturates disk but not CPU. How do you balance?
378. How do you tune JVM thread stack sizes?
379. What's your approach to tuning gossip threads?
380. How do you configure cross-DC workload replication balancing?
381. A cluster has uneven vnode ownership. How do you fix?
382. How do you tune Cassandra for mixed OLTP/OLAP workloads?
383. How do you configure thread pools for repair tasks?
384. A node consistently misses heartbeat checks. How do you tune?
385. How do you optimize streaming threads for new nodes?
386. What's your approach to handling thread pool backlogs?
387. How do you tune workload for IoT telemetry spikes?
388. A node drops messages under stress. How do you optimize?
389. How do you configure speculative\_retry for workloads?
390. How do you tune client-to-coordinator thread pools?
391. A query workload grows 10x. How do you rebalance?
392. How do you optimize workloads in multi-cloud deployments?
393. How do you tune GC threads in JVM for Cassandra?
394. What's your approach to memory balancing between caches?
395. A node experiences frequent timeouts. How do you balance load?
396. How do you configure driver-side load balancing?
397. How do you optimize workload balancing during schema migrations?
398. A repair job lags behind schedule. How do you tune threads?
399. How do you optimize streaming parallelism for large datasets?
400. How do you balance workloads across racks within a DC?

#### **Indexing, Optimizer Hints, Troubleshooting (401–500)**

401. How do you decide between SASI vs SAI indexes?
402. A secondary index causes slow queries. How do you fix?
403. How do you optimize queries with filtering on clustering columns?
404. What's your approach to optimizing full-text search in Cassandra?
405. How do you optimize workloads using SASI indexes for range queries?
406. A query fails due to too many tombstones. How do you troubleshoot?
407. How do you optimize materialized views in production?



408. What's the impact of multiple materialized views on write performance?
409. How do you tune compaction for index tables?
410. A SASI index query performs poorly. How do you debug?
411. How do you design indexes for high-cardinality fields?
412. What's your approach to optimizing queries with LIKE patterns?
413. How do you troubleshoot schema agreement failures?
414. A query uses ALLOW FILTERING on indexed columns. How do you optimize?
415. How do you optimize indexes for time-series workloads?
416. How do you troubleshoot failed repairs?
417. What's your approach to optimizing indexing for analytics?
418. A query times out due to coordinator overload. How do you debug?
419. How do you optimize SASI for geospatial queries?
420. How do you troubleshoot high GC pauses during queries?
421. A query always times out in one DC. How do you debug?
422. How do you optimize indexes for multi-tenant workloads?
423. What's the impact of indexing TTL columns?
424. How do you troubleshoot consistency level ALL failures?
425. A query must filter across multiple attributes. How do you index it?
426. How do you optimize SAI for text search workloads?
427. How do you debug failed streaming during node bootstrap?
428. What's your approach to optimizing joins at the application level?
429. How do you troubleshoot repair mismatches?
430. How do you optimize Cassandra for hybrid search workloads?
431. How do you design indexes for multi-region clusters?
432. A query causes coordinator OOM. How do you debug?
433. How do you optimize materialized views for heavy writes?
434. What's the role of index\_interval in performance tuning?
435. How do you troubleshoot high dropped messages during queries?
436. How do you optimize indexing for fraud detection use cases?
437. A query must fetch latest data always. How do you design indexes?
438. How do you troubleshoot schema migration issues?
439. How do you optimize queries mixing indexed + non-indexed columns?
440. How do you optimize SASI for prefix searches?
441. A cluster experiences high repair pressure. How do you fix?
442. How do you optimize query tracing in Cassandra?
443. What's your approach to hybrid indexes in Cassandra?
444. How do you debug client-side driver timeouts?
445. A query uses filtering on multiple columns. How do you index?
446. How do you optimize for search workloads using DSE Search?
447. A node crashes during index rebuild. How do you recover?
448. How do you tune queries using LIMIT with indexes?
449. How do you troubleshoot inconsistent read results?
450. What's your approach to optimizing large batch statements?

451. How do you optimize queries that must sort large result sets?
452. How do you troubleshoot read repair mismatches?
453. How do you design indexes for dynamic attributes?
454. A query triggers tombstone GC storms. How do you optimize?
455. How do you optimize SAI indexes for analytics queries?
456. How do you troubleshoot write timeout errors?
457. A query needs low latency at CL=LOCAL\_QUORUM. How do you tune?
458. How do you optimize materialized views for query flexibility?
459. How do you troubleshoot streaming failures during decommission?
460. A query must join multiple datasets. How do you optimize?
461. How do you optimize indexes for customer search apps?
462. How do you troubleshoot pending compaction backlogs?
463. How do you optimize for high write amplification workloads?
464. A query needs sub-second latency on billions of rows. How do you optimize?
465. How do you troubleshoot mismatched schemas across nodes?
466. How do you optimize Cassandra for hybrid OLTP/OLAP queries?
467. What's your approach to multi-dimensional indexing?
468. A query with ALLOW FILTERING is unavoidable. How do you optimize?
469. How do you troubleshoot frequent GC pressure?
470. How do you optimize indexes for logs search?
471. A query returns inconsistent results under CL=QUORUM. How do you debug?
472. How do you optimize materialized views for dashboards?
473. How do you troubleshoot slow schema agreement?
474. How do you optimize Cassandra for search-heavy e-commerce workloads?
475. What's your approach to handling indexing at scale?
476. A query mixes static + dynamic columns. How do you optimize?
477. How do you troubleshoot OOM during compaction?
478. How do you optimize Cassandra for personalization queries?
479. A query joins across multiple keyspaces. How do you handle?
480. How do you optimize SAI for real-time text search?
481. How do you troubleshoot consistency mismatch alerts?
482. How do you optimize indexing for IoT metadata?
483. A query must combine multiple partition keys. How do you optimize?
484. How do you optimize Cassandra for hybrid transactional + search workloads?
485. How do you troubleshoot high client connection timeouts?
486. A query retrieves large datasets repeatedly. How do you optimize?
487. How do you optimize indexes for fraud detection?
488. How do you troubleshoot compaction failure errors?
489. How do you optimize schema + indexes for geo-distributed workloads?
490. A query must filter on non-primary key attributes. How do you optimize?
491. How do you optimize Cassandra for real-time analytics?
492. How do you troubleshoot driver-side load balancing issues?
493. How do you optimize indexes for healthcare search workloads?

- 494. A query shows high variance in latency. How do you debug?
- 495. How do you optimize hybrid workloads with SAI indexes?
- 496. How do you troubleshoot massive tombstone warnings?
- 497. How do you optimize Cassandra for time-series analytics with indexes?
- 498. A query must support full-text search. How do you optimize?
- 499. How do you optimize Cassandra for high-throughput OLTP queries?
- 500. How do you troubleshoot and optimize coordinator-side bottlenecks?