Here are some Splunk query-related interview questions that you might be asked:

Basic Splunk Queries

1. How do you retrieve logs from a specific index and source in Splunk?

index=my_index source="/var/log/app.log"

2. How do you filter logs that contain a specific keyword like "error" or "failed"?

index=my_index "error" OR "failed"

3. How do you extract specific fields from log data?

Using rex:

index=my_index | rex "User: (?<username>\w+)"

Using spath for JSON logs:

index=my_index | spath input=_raw path=response.status

Aggregation & Statistical Functions

4. How do you find the number of errors per server?

index=my_index "error" | stats count by host

5. How do you calculate the average response time per application?

index=my_index | stats avg(response_time) by app_name

6. How do you find the top 5 most frequently occurring error messages?

index=my_index "error" | top limit=5 message

Time-Based Analysis

7. How do you create a time-based chart for CPU usage?

index=my_index sourcetype=cpu_usage | timechart avg(cpu) by host

8. How do you compare today's error count with yesterday's?

index=my_index earliest=-1d@d latest=@d "error"

| stats count as yesterday_count

| appendcols [search index=my_index earliest=@d "error" | stats count as today_count]

Performance & Optimization

9. How do you improve search performance in Splunk?

Use index and sourcetype filters.

Use tstats instead of stats where possible.

Avoid wildcard searches in large datasets.

Limit the number of fields retrieved using fields command.

10. What's the difference between stats and tstats?

stats operates on raw events, while tstats queries indexed data, making it faster.

Real-Time Monitoring & Alerts

11. How do you create an alert for high CPU usage (>90%)?

index=my_index sourcetype=cpu_usage | stats avg(cpu) as avg_cpu by host | where avg_cpu > 90

12. How do you detect unusual spikes in error rates?

index=my_index "error" | timechart count | anomalies

Complex Queries & Use Cases

13. How do you correlate user login failures with network issues?

index=auth_logs "login failed" | join user_id [search index=network_logs "connection timeout"]

14. How do you track the number of failed transactions per user within a 5-minute window?

index=my_index "transaction failed" | transaction user_id maxspan=5m | stats count by user_id

15. How do you extract IP addresses from logs?

index=my_index | rex field=_raw "(?<ip>\d+\.\d+\.\d+\.\d+)"

Scenario-Based Questions

16. If a client reports slow application performance, what Splunk queries would you run to investigate?

Check response times:

index=my_index sourcetype=app_logs | stats avg(response_time) by host

Check CPU/memory usage:

index=my_index sourcetype=cpu_usage OR sourcetype=memory_usage | stats avg(cpu) as avg_cpu, avg(memory) as avg_memory by host

17. If an application crashes at midnight daily, how would you identify the cause?

Search logs around that time:

index=my_index earliest=-1h@midnight latest=@midnight

Check service restarts:

index=my_index sourcetype=service_status status="restarted"

18. If disk space usage suddenly increases, how do you find the files responsible?

index=my_index sourcetype=disk_usage | stats sum(file_size) by file_path

Bonus: Challenge Questions

19. How do you find all unique error codes from logs?

index=my_index "error" | dedup error_code | table error_code

20. How do you create a Splunk query to monitor 95th percentile response time?

index=my_index | stats perc95(response_time)