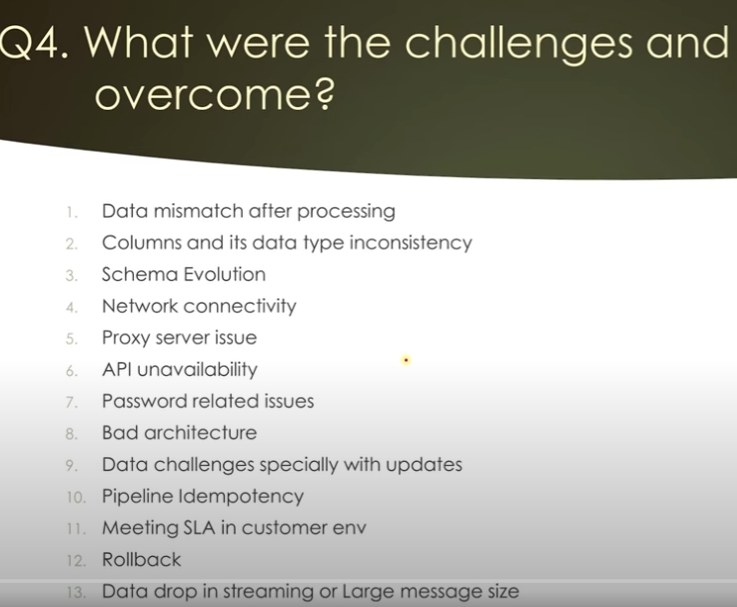


Explain end to end architecture



Connection pool in oracle when multiple sessions are open. How to use connection pool

Did you implement streaming pipeline

If my kafka consumer is reading slow then how will you handle back pressure

Kafka the definitive guide book

Youtube - ease with data by shubham

Learning journal - youtube

Why are you still not storing data in parquet

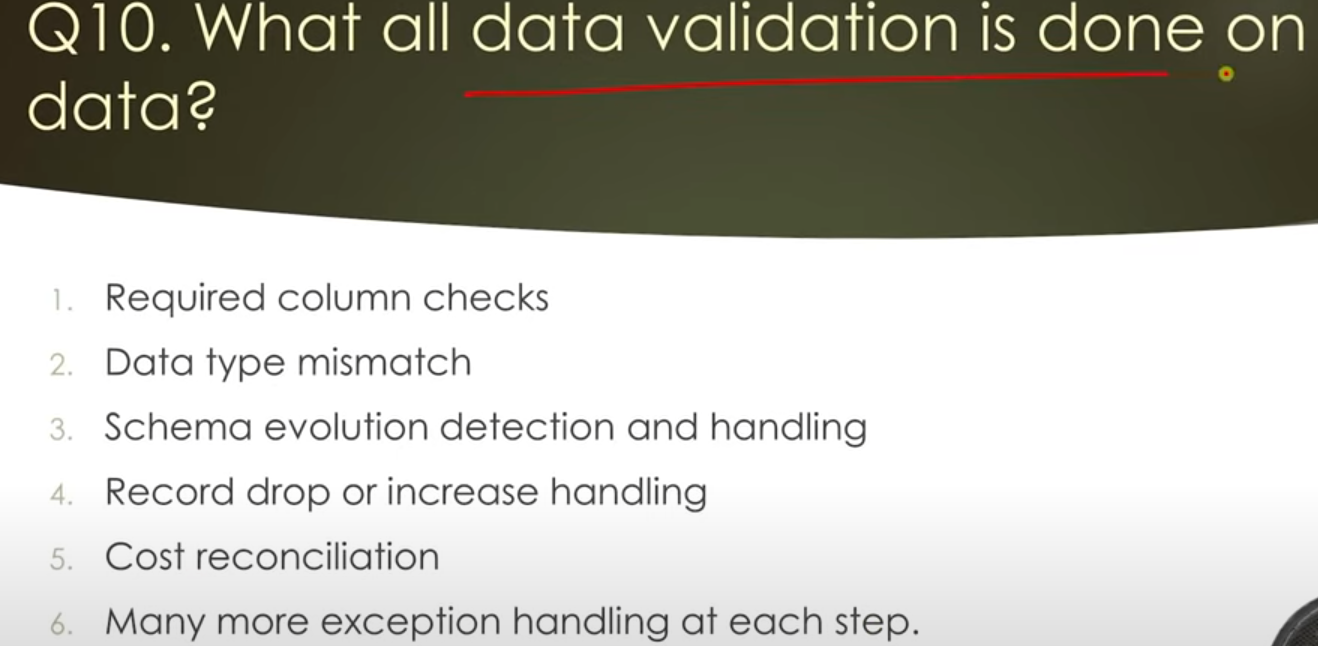
What is the volume of source data

2.3 million server’s records will process

5-6 CG

1TB monthly records

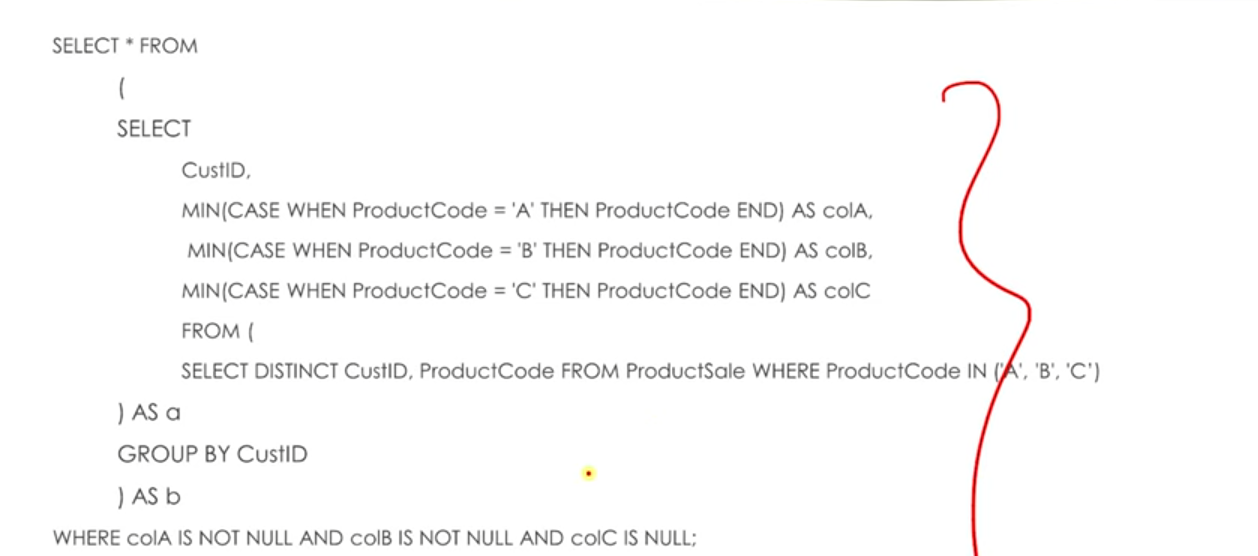
2yrs 12TB data

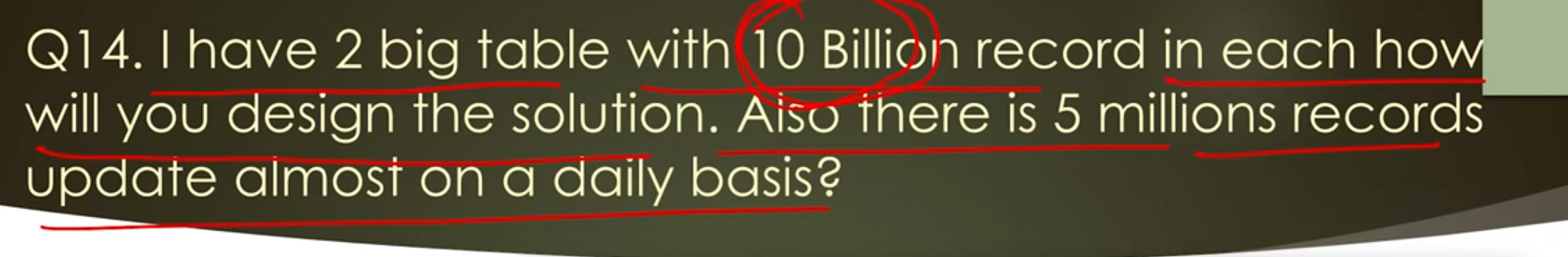


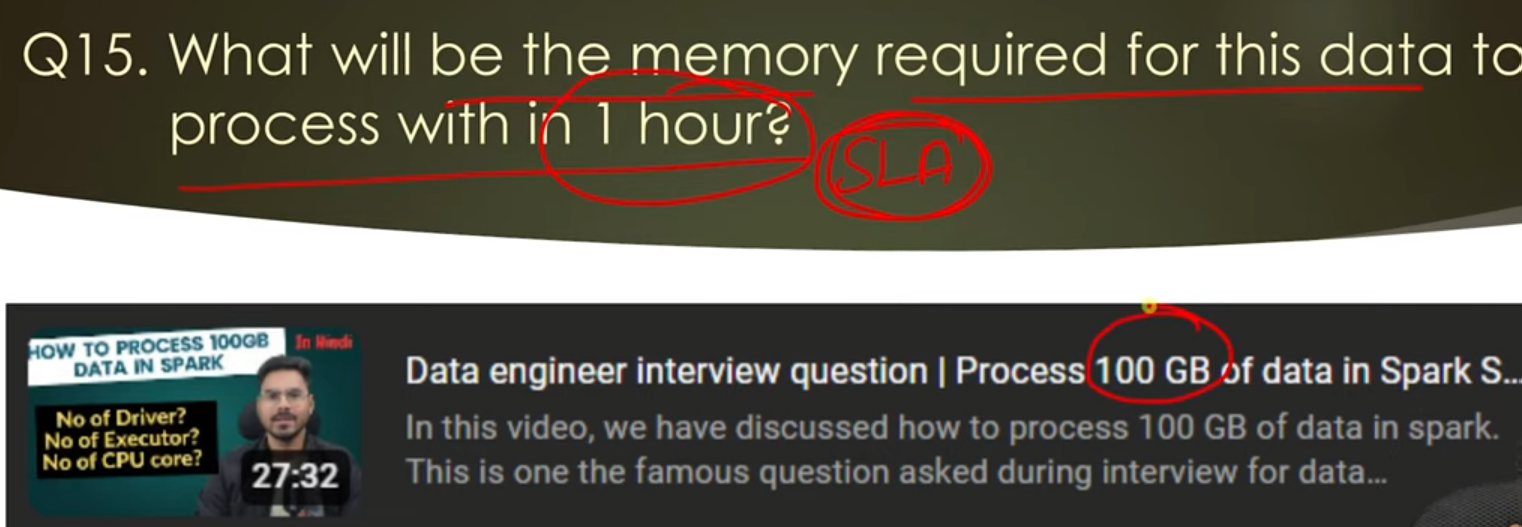
What stored procedure you use in your code

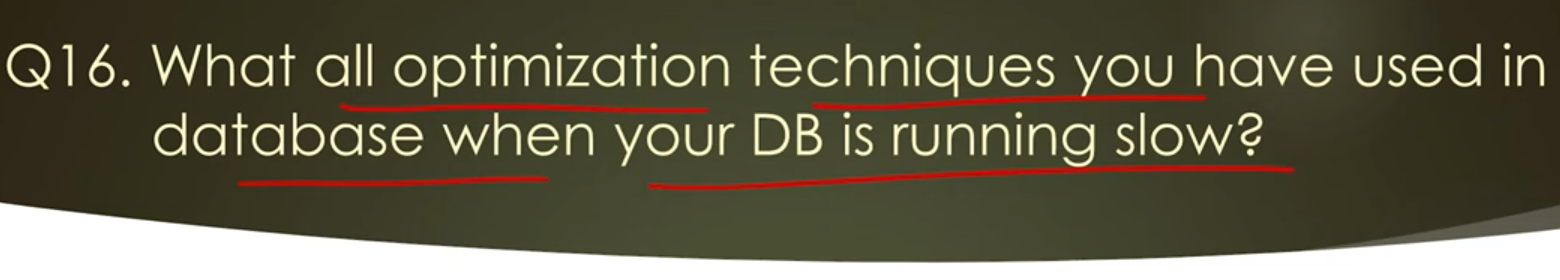


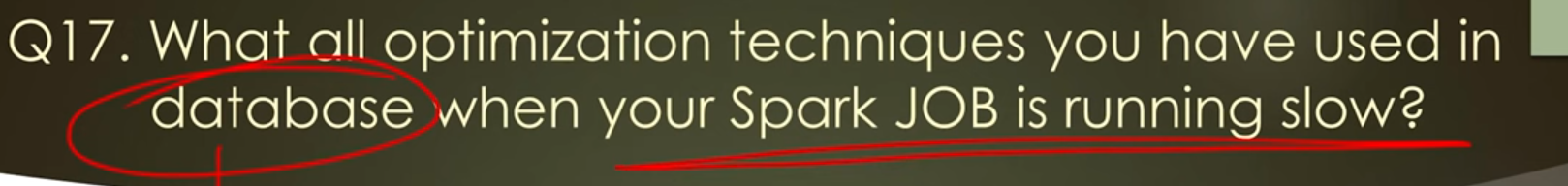
Each component why used, what is the alternative













How glue is different than EMR

What is system as a service

IAAS and PAAS difference

What microservices you have worked on (fast APIs knowledge is a plus)

Have you worked on ECS?

What all design pattern you have implemented in microservices

What is facade and circuit breaker design pattern

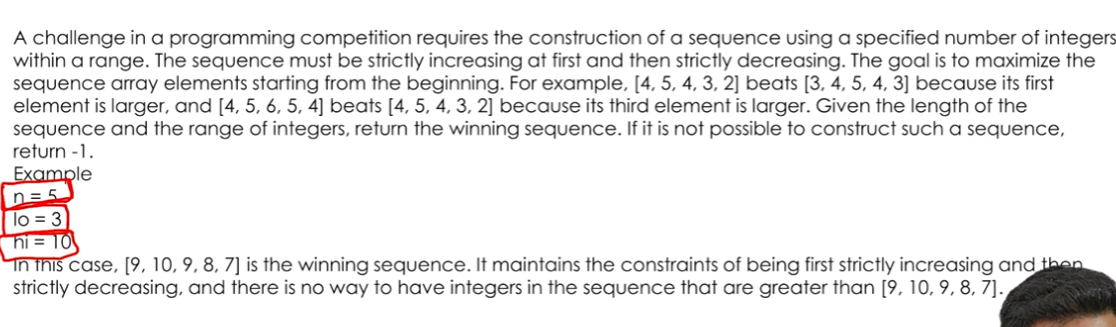
When do you use coalesce

What is the role of partitioning and when writing the data in parquet, how do you decide partitions columns

What is parquet file & why it is faster

What is incremental load

how checkpointing works in streaming



System design to build streaming solution

If you have disagreement with your manager, what will you do (behavioural questions)

Your teammate is lagging in work, what will you do

How do you implement end to end SDLC process?

How do you implement data quality in your pipeline

What are the measures you take so that your project does not brake in production

How do you roll the feature to first few % age of customer to make sure whether feature is working properly or not