Chapters:

00:00 - Introduction

01:45 - Architecture of Spark

03:35 - How Spark is an in-memory computing Engine?

06:15 - What is the difference between RDD, Dataframe and dataset and why do we use them?

07:28 - What are cache and persist and what is the difference between them?

09:30 - Spark Optimization

12:23 - Shuffling in Spark

14:18 - Examples of Wide and Narrow Transformation

15:00 - Why do we have GroupByKey, ReduceByKey, SortByKey?

15:50 - Spark Scenario Based Question - 1

30:59 - Trade-off between Spark SQL and DataFrame

32:41 - DSA Question - 1