How Apache handled data skew in Spark 3?

Swipe ←

What is Data Skew?

When there is a sudden increase in data into one partition over the others, that is data skewness.

Impact of Data Skew?

Executor with big Partition takes longer time to complete the task while other executors take very little time. It beats the concept of parallel processing as one executor is performing the majority of work (sometimes 99%) whereas others are just performing very little work(sometimes just 1%).

Why Data Skew happens?

1. Unevenly distribution of data is one of the major reasons.
2. Poor programming and architecture can also lead to data skewness.
3. Sometimes Coalesce could also be a culprit .

How do Data Engineers deal with it before Spark 3?

1. It requires a lot of data analysis and makes changes in code accordingly. Constantly monitor Spark UI for calculation and stats.
2. Salting is one of the famous solutions for this.
3. Repartition, but it could be more expensive sometimes.

What Spark3 has for data skewness?

Adaptive Query Execution (AQE).

It is an optimization technique in Spark SQL that makes use of the runtime statistics to choose the most efficient query execution plan (definition taken from the spark document). It eliminates the need to manually tracking the Spark UI stats.

AQE has 3 Features

Coalesce Post Shuffle Partitions

Dynamically Switch Join Strategies

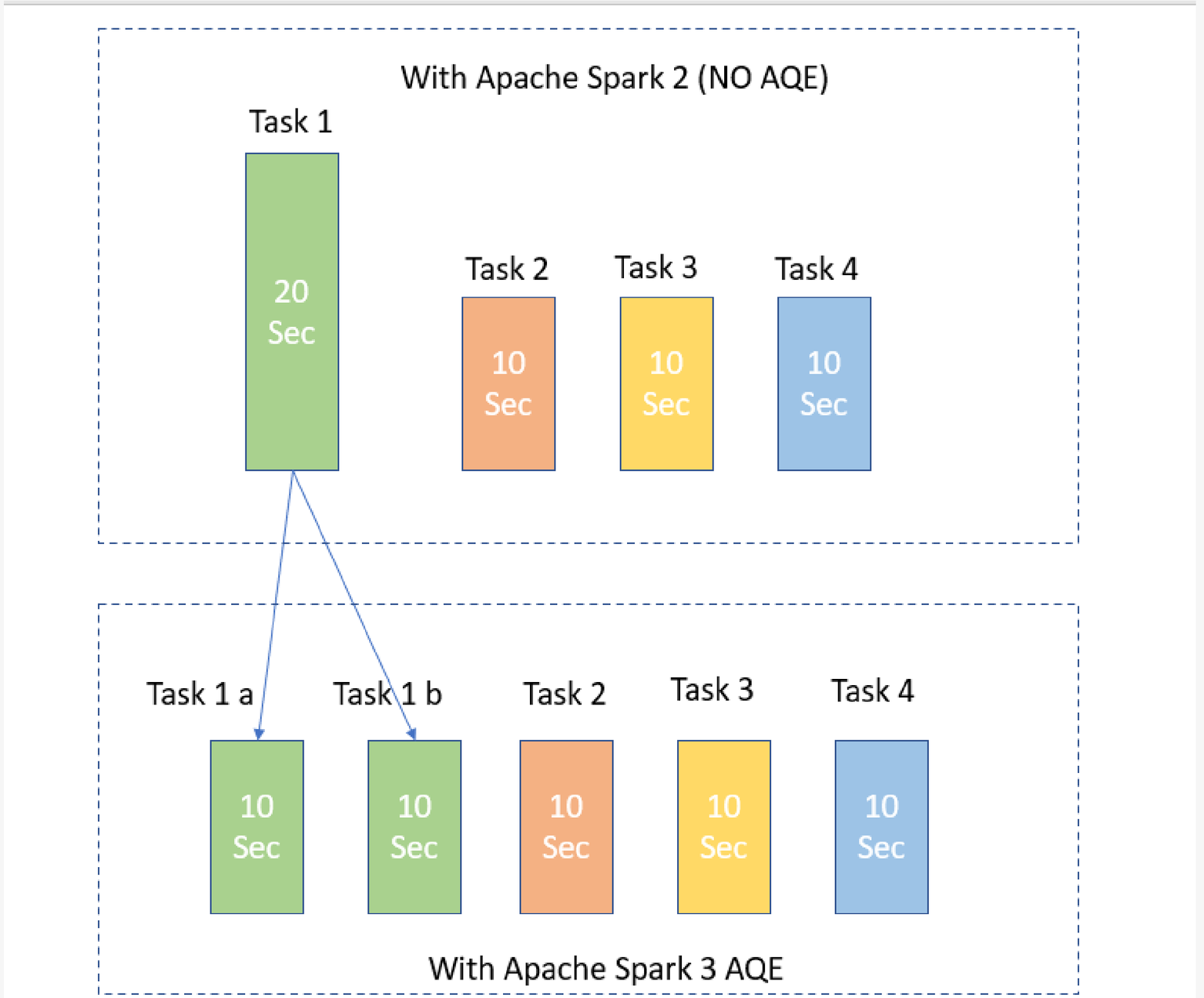
Dynamically Optimized Skew join

Dynamically optimizing Skew Joins

AQE detects the skew join automatically from the shuffle statistics, which will then divide a long time taking tasks into smaller tasks. Hence it increases the speed of overall job.

Example

Consider Stage 1 has 4 tasks: task 1 takes 20 sec to complete where task 2,3,4 takes 10 seconds each. Spark will divide Task 1 into two subtasks which will complete the subtask in 10 seconds each.


With Apache Spark 2 (NO AQE)
Task 1
20 Sec
Task 2
10 Sec
Task 3
10 Sec
Task 4
10 Sec
Task 1 a
10 Sec
Task 1 b
10 Sec
Task 2
10 Sec
Task 3
10 Sec
Task 4
10 Sec
With Apache Spark 3 AQE

Thanks for Reading, Follow for more such posts.