

Streaming Joins



Objective

Join data in micro-batches

Structured Streaming join API



Takeaways

Same join API as non-streaming DFs

```
val joinCondition = guitaristsDF.col("band") === bandsDF.col("id")
val guitaristsBandsDF = guitaristsDF.join(bandsDF, joinCondition, "inner")
```

Some join types are not supported

- RIGHT outer and RIGHT semi joins with static DFs
- LEFT outer and LEFT semi if static DF is joining against streaming DF
- full outer joins

Stream-stream joins are supported

- inner join optionally with a watermark*
- left/right outer join ONLY with a watermark*

Spark rocks

