

<https://databricks.com/session/deep-dive-into-project-tungsten-bringing-spark-closer-to-bare-metal>

https://spoddutur.github.io/spark-notes/deep_dive_into_storage_formats.html

<https://databricks.com/blog/2016/07/14/a-tale-of-three-apache-spark-apis-rdds-dataframes-and-datasets.html>

<https://www.appsflyer.com/blog/the-bleeding-edge-spark-parquet-and-s3/>

<https://stackoverflow.com/questions/34382612/how-to-avoid-generating-crc-files-and-success-files-while-saving-a-dataframe>

<https://stackoverflow.com/questions/35002184/spark-jobs-finishes-but-application-takes-time-to-close>

<https://medium.com/quantumblack/spark-udf-deep-insights-in-performance-foa95a4d8c62>

<https://tukaani.org/lzma/benchmarks.html#:~:text=Conclusions-,Compression,15%25%20smaller%20files%20than%20gzip.>