

Spark Performance

Techniques for Improving Spark Performance

- ▶ Tuning Spark configuration to improve scaling for large workloads, enhance parallelism, and minimize memory starvation among Spark executors.
- ▶ Using caching and persistence with appropriate levels to expedite access to frequently used data sets.
- ▶ Using the Spark web UI to gain visual perspective on performance.
- ▶ Using the right data format.