

Click to prove
you're human



Since it could potentially take seconds for a large shuffle to finalize, having multiple threads helps driver to handle concurrent shuffle merge finalize requests when push-based shuffle is enabled. 3.3.0 spark.shuffle.push.minShuffleSizeToWait 500m Driver will wait for merge finalization to complete only if total shuffle data size is more than this threshold. If total shuffle size is less, driver will immediately finalize the shuffle output. 3.3.0 spark.shuffle.push.minCompletedPushRatio 1.0 Fraction of minimum map partitions that should be push complete before driver starts shuffle merge finalization during push based shuffle. 3.3.0