



Figure 1: A comparison of time to complete tasks from measurements reported in the Spark event log. The difference in the reported launch and finish times, $t_2 - t_1$, is compared to the sum of the elapsed times reported for: serialization, garbage collection, and execution components of each task.