Abstract

Scheduling Hadoop is a way to set up any job that will run on hadoop system that can manage all the jobs to get a turn in the execution on every resource available. Default Scheduling in Hadoop, FIFO that has characteristics first come first excecution and monopoly of the entire resource. But FIFO has disadvantage for short job that run into starvation when a long job in execution.

Fair share improve Delay Scheduling is a job scheduler that use a methode to divide job for one cluster in some pool. For each pool that has been divided, put in place a method to Delay the course of the next jobs to improve data locality before. Resource sharing of data and the data that is nearly optimal allocation will affect the Job Fail Rate, Job Throughput, Average Completion Time. Delay improve Fair share have effective performance than Fair share and Delay Scheduling in the kind of job randomtextwriter in data type .txt with a reduction of fail rate 0,3% job and increase produce of throughput 2,73 job/m and faster 273,59 minutes than Delay Scheduling and 128,15 minutes than Fair share.

Keywords: hadoop, multi-node hadoop, Fair share scheduling Delay improve improve capacity, improve capacity scheduling Delay