Scheduling divisible jobs to optimize the computation and energy costs

ABSTRACT

The important challenge in cloud computing environment is to design a scheduling strategy to handle jobs, and to process them in a heterogeneous environment with shared data centers. In this paper, we attempt to investigate a new analytical framework model that enables an existing private cloud data-center for scheduling jobs and minimizing the overall computation and energy cost together. Our model is based on Divisible Load Theory (DLT) model to derive closed-form solution for the load fractions to be assigned to each machines considering computation and energy cost. Our analysis also attempts to schedule the jobs such a way that cloud provider can gain maximum benefit for his service and Quality of Service (QoS) requirement user's job. Finally, we quantify the performance of the strategies via rigorous simulation studies.

Keyword: Cloud computing; Divisible load theory; Energy saving; Scheduling; QoS