Predicting Data Center Resource Usage Using Quantile Regression to Conserve Energy While Fulfilling the Service Level Agreement

Authors : Ahmed I. Alutabi, Naghmeh Dezhabad, Sudhakar Ganti

Abstract : Data centers have been growing in size and dema nd continuously in the last two decades. Planning for the deployment of resources has been shallow and always resorted to over-provisioning. Data center operators try to maximize the availability of their services by allocating multiple of the needed resources. One resource that has been wasted, with little thought, has been energy. In recent years, programmable resource allocation has paved the way to allow for more efficient and robust data centers. In this work, we examine the predictability of resource usage in a data center environment. We use a number of models that cover a wide spectrum of machine learning categories. Then we establish a framework to guarantee the client service level agreement (SLA). Our results show that using prediction can cut energy loss by up to 55%.

Keywords : machine learning, artificial intelligence, prediction, data center, resource allocation, green computing
Conference Title : ICGCTI 2022 : International Conference on Green Computing, Technology and Innovation
Conference Location : Berlin, Germany
Conference Dates : May 23-24, 2022