Abstract

Cloud computing emerged as an efficient alternative to physical servers and provided benefits such as elasticity and scalability to the end users. Virtualization allowed dynamic allocation and deallocation of resources in cloud computing data centers and hence a very convenient model based on pay-as-you-go and optimal resource utilization became immensely popular. The growth of cloud popularity and the reliance of various applications on cloud was however unprecedented and no one could possibly predict that a novel paradigm would so quickly become so heavily used. Areas like Body Area Networks, Intra-vehicular communication, smart cities and smart buildings and smart traffic signaling etc. made cloud computing so popular that it is now predicted that the number of end users of the cloud networks would be around four times the population of Earth by year 2050 AD. In wake of this advent of IOT and the evident shortcomings of cloud networks, a new layer of processing, compute and storage was introduced by CISCO in 2014, namely the Fog Layer. This thesis is aimed at efficiently distributing workloads between the Fog Layer and the Cloud Network and then optimizing resource allocation in cloud networks to ensure better utilization and quick response time of the resources available to the end user. We have employed a Dead-line aware scheme to migrate data between cloud and Fog networks based on data profiling and then used K-mean clustering and service request prediction model to allocate resources efficiently to all requests. To substantiate our model we have used iFogSim which is an extension of the CloudSim simulator.