Abstract

Cloud computing refers to the model, which is the pool of resources. Cloud makes on-demand delivery of these computational resources (data, software and infrastructure) among multiple services via a computer network with different load conditions of the cloud network. User will be charged for the resources used based upon time. Hence efficient utilization of cloud resources
has become a major challenge in satisfying the user’s requirement (QoS) and in gaining benefit for both the user and the service provider. In this paper, we propose a priority and admission control based service scheduling policy that aims at serving the user requests satisfying the QoS, optimizing the time the service-request spends in the queue and achieving the high throughput of the cloud by making an efficient provision of cloud resources.

Reference

- Cloud computing - Wikipedia, the free encyclopedia.htm
- Chen Ming1 Li Mengkun1, Cai Fuqin1 “A model of scheduling optimizing for cloud computing resource services based on”, 2010 IEEE International conference on Granular Computing. DOI 10-1109/GiC 2010.180

Index Terms

Computer Science

Pattern Recognition

Key words

Subscription-category

service scheduling policy
Priority

admission control and deadline