CoolCloud: Improving Energy Efficiency in Virtualized Data Centers

Abstract:

Recent years, cloud computing services continue to grow and become more pervasive and indispensable in people's lives. The energy consumption continues to rise as more and more data centers are being built. How to design a more energy efficient data center has become one of the most important issues for cloud computing service providers.

In this thesis, we propose an optimization framework to minimize energy consumption in virtualized data centers. The proposed framework minimizes energy at two different layers: (1) minimize local server energy using dynamic voltage & frequency scaling (DVFS) exploiting runtime program phases. (2) minimize global cluster energy using dynamic mapping between virtual machines (VMs) and servers based on each VM's resource requirement. Such optimization leads to the most economical way to operate an enterprise data center.

On each local server, we develop a voltage and frequency scheduler (we call it the "cool scheduler") that can provide CPU energy savings under applications' specified SLA requirements by exploiting applications' run-time program phases. This design greatly improves the computation efficiency compared to other published works. The scheduler contains a feedback design to handle frequency mismatch for each thread in order to maintain the application service level agreement. The scheduler also minimizes the overhead of DVFS operations by grouping threads that has the same frequency need onto the same CPU core using thread migration. At the cluster level, we propose a practical solution for managing the mappings of VMs to physical servers. This framework solves the problem of finding the most energy efficient way (least resource wastage and least power consumption) of placing the VMs considering their resource requirements. Due to that the optimization is NP hard, a heuristic approach is further proposed to reduce computation complexity and make our design scale well to the size of enterprise data centers.