

Towards Scalable Traffic Management in Cloud Data Centers

Abstract:

Cloud Computing is becoming a mainstream paradigm, as organizations, large and small, begin to harness its benefits. This novel technology brings new challenges, mostly in the protocols that govern its underlying infrastructure. Traffic engineering in cloud data centers is one of these challenges that has attracted attention from the research community, particularly since the legacy protocols employed in data centers offer limited and unscalable traffic management. Many advocated for the use of VLANs as a way to provide scalable traffic management, however, finding the optimal traffic split between VLANs is the well known NP-Complete VLAN assignment problem. The size of the search space of the VLAN assignment problem is huge, even for small size networks. This paper introduce a novel decomposition approach to solve the VLAN mapping problem in cloud data centers through column generation. Column generation is an effective technique that is proven to reach optimality by exploring only a small subset of the search space. We introduce both an exact and a semi-heuristic decomposition with the objective to achieve load balancing by minimizing the maximum link load in the network. Our numerical results have shown that our approach explores less than 1% of the available search space, with an optimality gap of at most 4%. We have also compared and assessed the performance of our decomposition model and state of the art protocols in traffic engineering. This comparative analysis proves that our model attains encouraging gain over its peers.