Distributed Server Migration for Scalable Internet Service Deployment

Abstract:

The effectiveness of service provisioning in large-scale networks is highly dependent on the number and location of service facilities deployed at various hosts. The classical, centralized approach to determining the latter would amount to formulating and solving the uncapacitated Median (UKM) problem (if the requested number of facilities is fixed) or the uncapacitated facility location (UFL) problem (if the number of facilities is also to be optimized). Clearly, such centralized approaches require knowledge of global topological and demand information, and thus do not scale and are not practical for large networks. The key question posed and answered in this paper is the following: "How can we determine in a distributed and scalable manner the number and location of service facilities?" In this paper, we develop a scalable and distributed approach that answers our key question through an iterative reoptimization of the location and the number of facilities within network neighborhoods. We propose an innovative approach to migrate, add, or remove servers within limited-scope networkneighborhoods by utilizing only local information about the topology and demand. We show that even with limited information about the network topology and demand, within one or two hops, our distributed approach achieves performance, under various synthetic and real Internet topologies and workloads, that is comparable to that of optimal, centralized approaches requiring full topology and demand information. We also show that it is responsive to volatile demand. Our approach leverages recent advances in virtualization technology toward an automated placement of services on the Internet.