ABSTRACT

The Internet has evolved to become a critical infrastructure for content delivery. Typical customers of a Content Distribution Network (CDN) are media and Internet advertisement companies, data centers, Internet Service Providers (ISPs), online music retailers, mobile operators, consumer electronics manufacturers, and other carrier companies. Each of these customers want to publish and deliver their content to the end-users on the Internet in a reliable and timely manner. However, the Internet being a large, highly-dynamic, heterogeneous and untrusted environment raises numerous challenges for building efficient and fault tolerant content delivery infrastructure. To move beyond these shortcomings, this thesis establishes the basis for developing an efficient and fault tolerant content delivery solutions that are efficient, high performance, and fault-tolerant.

The main goal of the Content Distribution Network (CDN) is to replicate the content from the origin server to surrogate servers, scattered over the globe, in order to deliver the content to end users in a reliable and timely manner from nearby surrogate servers. The performance of content distribution network is analyzed to ensure that whether the user perceived Quality of Service (QoS) in content delivery is maintained. Quality of Service in content delivery is ensured by minimizing the impact of surrogate server failure as well as overcoming the overload problem that is a serious threat for busy sites serving popular contents.
The main objective of this thesis is the application of the dominating set and its variants to construct the semantic overlay network of surrogate servers which in turn provides the logical infrastructure of the content distribution network. Overlay networks play an important role in group communication applications in the Internet. It effectively uses the Internet as a low level infrastructure, to provide distributed content delivery services to end users.

First we propose a novel Efficient and Fault Resilient Replication Algorithm (EFRRA) to disseminate the content among the surrogate servers in a reliable and timely manner. This EFRRA algorithm is able to replicate the content among the surrogate servers and also capable of providing fault resilient property during surrogate server failure. EFRRA’s performance is evaluated in terms of Average Replication Time and Maximum Replication Time.

In this thesis, for designing an efficient and resilient content distribution network, the following algorithms are proposed to construct the logical infrastructure of the semantic overlay network.

- Dominating set formation algorithm
- Equitable Dominating set formation algorithm
- QoS aware Dominating set formation algorithm

In dominating set based semantic overlay network construction algorithm, the optimum number of dominant surrogate servers are grouped
and logically connected by finding the sub graph of the original network. It is also ensured that all the surrogate servers are either member of dominant set or adjacent set.

In equitable dominating set based semantic overlay network construction algorithm, the optimum number of dominant surrogate servers are grouped such that the difference between the degrees of all the surrogate servers in the dominant set of surrogate servers can differ utmost by 1 and logically connected by finding the sub graph of the original network. Alternatively, we described that each surrogate server in equitable dominating set has more or less same number of one hop neighbor surrogate servers which are members of adjacent set.

The core of this thesis is constructing a QoS aware dominating set based semantic overlay network which satisfies the user perceived QoS constraints such as efficiency, controlled redundancy, and fault tolerant. The performance results of the proposed QoS aware Dominating set based semantic overlay network for content distribution is analyzed in terms of Size of CDN, Mean Response Time, Latency, Hit Ratio Percentage, Number of Completed Requests, Rejection rate and Mean CDN load.