ABSTRACT

The growing ubiquity of the Internet is changing the nature of software design and deployment. Internet-based system architectures employ distributed components and use mobile code to deliver new software to millions of users. By a lack of understanding of the processor (CPU) time required by such dynamically injected software, computer operating systems cannot effectively manage system resources or control the execution of the mobile code. There is a need to quantify CPU time requirements of the mobile code in a form that can be understood readily on heterogeneous computing platforms.

An extensive research is being carried out on active networks, an emerging technology that exploits the extreme features of the mobile code. In an active network, network nodes in addition to forwarding messages, perform customized computation on messages flowing through them. While active networks offer flexibility in tailoring network services to distributed applications, one of the pitfalls of using active networks is the potential degradation in overall application performance when multiple applications use the active node services simultaneously. The inherent unpredictability of the execution time of an arbitrary execution code in the active packet poses a significant challenge in CPU scheduling and in providing QoS guarantees for data flows that compete for resources in the network. The resources in active networks comprise both the CPU and bandwidth. The resource allocation
techniques used for traditional networks do not directly extend to active networks. Efficient estimation methodologies and allocation policies are required for the optimal utilization of the resources in active networks.

The main objective of this work is to study the CPU control and estimation methods for efficient resource control and to design an algorithm for fair resource control and allocation in active networks based on the resource estimation for providing Quality of service guarantees to the competing flows. The estimation techniques are analyzed for predicting the processing requirements prior to allocation. The performance of these prediction techniques are evaluated and compared against the actual allocation. The scheduling algorithm that is developed satisfies the fair resource allocation constraint by adjusting the share of CPU resource given to each flow based on the share of bandwidth given to the flow.

Another objective of the work is to perform resource prediction and allocation using policies as a Knowledge Based Expert System (KBES). Policy based allocation of resources can provide a predictable dynamic environment. So the objective is to design an algorithm based on policies and priorities and to analyze the performance with and without priorities. A study of active networks with security issues, denial of service attacks, and distributed denial of service attacks is conducted as a case study. The performance of various queuing algorithms is examined for efficient resource allocation during distributed denial of service attacks that occur in an active network environment.
Distributed applications such as online stock quote, online auction, and web caching and fusion applications can exploit the services offered by the active network to improve its performance as well. The similarities between active applications in active network execution environments and other forms of dynamically injected softwares in Internet based software architectures encourage us to believe that the estimation and allocation methods apply generally to the problem of specifying resource demand in distributed applications that rely on the use of the mobile code.