Cluster-based Web Services using an Integrated Adaptive Management System

Prof. Kulvinder Singh Mann
Assistant Professor in CSE/IT Deptt.
Guru Nanak Dev Engg. College
Ludhiana.

ABSTRACT

A cluster based web server is a web system architecture that consists of multiple server nodes distributed on a local area and is interconnected through a high speed network. The complexity of the cluster-based web service challenges the traditional approaches, which fail to guarantee the reliability and real-time performance required. In this paper, we present an Integrated Adaptive Management System (IAMS) for such service.

The issues we discuss address to efficiently allocate resources and provide more effective QoS support under a wide range of load conditions. For the global resource level, we introduce spare instance and corresponding management strategy as a supplemental adaptive mechanism. The spare instances hosted on shared node afford better resource utilization and more effective QoS support in the case of overload or workload fluctuation. Further, it can relax the influence of the fault recovery from the hardware and software failure. For the local level, we apply a multipurpose linear-quadratic regulator (LQR) as basic adaptive element. The control scheme using reject time ratio as control input is able to provide guarantees for overload protection, resource control, Qos control, performance isolation, and effective management for spare instances. Although cluster-based web services have been widely deployed we have seen limited research in the literature on comprehensive adaptive mechanism for resource management with QoS support. For lack of detailed knowledge about the inner workings of the web service, we adopt system identification approach to establish the models. The open-loop service is modeled as a difference equation with unknown parameters. Then we stimulate the web service with pseudo-random signal as inputs and sample output data. After collecting the data, the least square estimation of model parameters can be solved via MATLAB.

In this paper, we present an integrated adaptive management system (IAMS) for cluster-based web services. The main adaptive mechanisms in our system are: The resource adaptation based on the SLA event-driven mechanism. A supplemental adaptive mechanism based on spare instances and the corresponding strategy. A multi-purpose control scheme incorporated in the local manager and/or application is used as the basic adaptive element to provide guarantees for overload protection, resource control, Qos control, performance isolation and management for spare instances. We have presented rich experiments including overload protection, resource utility and QoS performance control both on static and dynamic website. The results illustrate our multi-purpose control scheme can offer effective QoS differential service and satisfied performance. Now the controller as the adaptive element only provides throughput different-tiation. It discusses the spare instance and its management. The multi-purpose feedback control scheme and the implementation of QoS differentiated service. It presents the experiment results on Tomcat and a 3-tiered web site to illustrate the efficiency of our control scheme respectively. Last Section summarizes our work and discusses the future work.

Results of experiments on both static and dynamic web sites illustrate the efficiency and robustness of the multi-purpose LQR.