Chapter 9

Peer-to-Peer Platforms for High-Quality Web Services: The Case for Load-Balanced Clustered Peer-to-Peer Systems

Ying Qiao
University of Ottawa, Canada

Shah Asaduzzaman
University of Ottawa, Canada

Gregor V. Bochmann
University of Ottawa, Canada

ABSTRACT

This chapter presents a clustered peer-to-peer system as a resource organization structure for web-service hosting platforms. Where service quality, such as response time and service availability, are provided with assurance. The peer-to-peer organization allows integration of autonomous resources into a single platform in a scalable manner. In clustered peer-to-peer systems, nodes are organized into clusters based on some proximity metric, and a distributed hash table overlay is created among the clusters. This organization enables lightweight techniques for load balancing among different clusters, which is found to be essential for providing response time guarantees. Service availability is provided by replicating a service instance in multiple nodes in a cluster. A decentralized load balancing technique called diffusive load balancing is presented in the context of clustered peer-to-peer systems and evaluated for effectiveness and performance.

INTRODUCTION

Web services are autonomous software systems that can be advertised, discovered and accessed through exchanging messages over the Internet using a set of standard protocols (e.g. SOAP, WSDL, UDDI). The standard protocols bring the interoperability among these autonomous software systems and allow creation of more complex and powerful applications through web service composition. This naturally allows distributed computation through execution of different components of an application.
hosted at autonomous Internet hosts. Computing resources, such as data, storage and CPU processing power of the hosts are thus shared by different users across the Internet.

When web services technology makes it easy to have distributed computing among computers and applications with different platforms, architectures, and programming languages, the role of distributed computing has expanded from assisting daily routines inside an enterprise to participating in the interactions among enterprises. However, the challenge to realize web service applications in a large scale still remains.

As web services are provided by software executed in computers on the Internet, situations like overloaded or failed servers or congested networks can largely affect the quality of web services, where clients could experience unexpected delay and jitter on accessing web services. Consequently, the quality of a web service application can hardly be guaranteed with the unequal performance of its services. This directly challenges the further development of web service applications.

This situation becomes more and more severe with the large deployment of web service applications; it requires a scalable and efficient execution platform to provide high quality services. This chapter discusses the usability of a clustered peer-to-peer system with explicit load-balancing schemes as a platform for hosting web services. Web services hosted on server-pool based hosting platforms or data centers often suffer from overloaded servers and congested networks due to the static scale of the platform, which seriously degrades the perceived performance of the services. Highly scalable and adaptable peer-to-peer computing platforms may be desirable in these scenarios. The capability of peer-to-peer systems to efficiently organize large number of Internet-connected computers without any centralized controller makes it a good candidate as a platform for web services. Although there have been several works proposing the use of peer-to-peer techniques for web services discovery (Verma, 2005; Schmidt, 2004), attempts to exploit the scalability of peer-to-peer computing platforms for quality assured hosting of web services are limited.

Peer-to-peer computing platforms are characterized by a huge collection of autonomous and inconsistent resources. A characteristic behavior of the resources in a peer-to-peer system is their intermittent arrival and departure, which is called churn. Explicit techniques are necessary to provide reliable and homogeneous services abstracting this behavior. Studies show that organizing these autonomous resources in clusters improves the reliability and the robustness of the platform (Locher, 2006). Clustered peer-to-peer systems such as eQuus improves the robustness of the system against churn by organizing the resources in clusters and replicating the resource-states within a cluster; applications built over these systems could have a more deterministic performance, e.g., CliqueStream (Asaduzzaman, 2008) for delivering life video streams over eQuus.

For the heterogeneity of the autonomous resources and the uneven distribution of service requests, some explicit load balancing mechanism is required to smooth out the performance inequalities towards some assured level of quality. A load balancing system can arrange resources according to the state of computing nodes and requests, so that the overall performance of the system can be improved. After using load balancing techniques, different web services are expected to have consistent performance characteristics across the platform.

Here we propose a hosting platform for web services using a load-balanced clustered peer-to-peer system. The platform performs load balancing at two levels: intra-cluster, i.e., loads among nodes in a given cluster are balanced, and inter-cluster, i.e., loads among different clusters are balanced. Intra-cluster load balancing is achieved through commonly used techniques such as request routing, and the inter-cluster load balancing is achieved through migration of resources between different clusters. Organization of the physical resources in