Chapter 59
Resource Scheduling Techniques in Utility Computing: A Survey

Inderveer Chana
Thapar University, India

Tarandeep Kaur
Thapar University, India

ABSTRACT

Utility Computing offers on-demand services from a shared pool of resources and can be envisaged to be a benchmark in the IT development. The capability to provide on-demand services involves management of large number of resources that are geographically dispersed and thus poses a number of resource management and scheduling challenges in the domain of resource heterogeneity, dynamic resource locations and load balancing. Proficient resource allocations and efficient scheduling helps in achieving optimal resource utilization and hence enhances the performance of the system. This paper evaluates existing resource management systems, listing their key characteristic features and highlighting the factors that make the existing systems excel upon each other. It also discusses various resource scheduling techniques currently available and characterizes the techniques based on Quality of Service (QoS) parameters supported by them along with the classification on basis of their operating environment and further extends towards load balancing and energy efficiency support if available.

INTRODUCTION

Utility computing model depicts how computing needs of the users can be fulfilled in future IT industry. The underlying design of utility computing is based on a service provisioning model, where the users (consumers) pay the service providers for using computing power only when they need. Its analogy is derived from the real world where service providers maintain and supply utility services to the users, such as electrical power, gas, and water to the consumers who in turn pay

DOI: 10.4018/978-1-4666-9562-7.ch059

service providers based on their usage (Yeo et al., 2006). The growth of Utility computing appears like realization of the vision of Leonard Kleinrock, who believed that as computer networks grow up and become sophisticated, there will be spread of 'computer utilities', which, like present electric and telephone utilities, will service individual homes and offices across the country (Yeo et al., 2006).

The utility computing model offers a number of benefits to both service providers and users. From the provider's perspective, physical hardware and software components are not set up or configured. Instead, virtualized resources are created and dynamically allotted to various users requests. The providers can thus easily and quickly reallocate resources with highest demands (Wu & Buyya, 2010). In turn, this efficient resource allocation helps to minimize the operational costs for providers. Utility computing enables providers to achieve a better Return on Investment (ROI) (Wu & Buyya, 2010).

For users, the most prominent advantage of utility computing is the reduction of IT-related functional costs and associated complexities. Users no longer have to invest heavily or undergo any difficulty in building and maintaining the IT infrastructures with no concern for the possible over or under-utilization of IT infrastructures. This outsourcing model thus provides increased flexibility and ease for users to adapt to their changing business needs and environments (Wu & Buyya, 2010).

RESEARCH METHOD

Research Questions

This paper considers research perspective of the resource scheduling in Utility computing systems. It aims at summarizing the current state of the art in Utility resource scheduling research by proposing answers to the following questions:

1. What is the role of resource allocation and scheduling in Utility computing systems?
2. Which resource management systems are available in the market and what features make them excel over each other?
3. What resource scheduling methods are available and how much they ensure optimal resource utilization in addition to achieving economies of scale?
4. Which scheduling strategy best optimizes program execution and usage of necessary infrastructures?
5. Which resource scheduling strategy can cater to the changing market demands and provide the clients with the QoS they require?
6. Whether any scheduling technique supports balancing of loads among resources, minimizing energy wastage?

Study Selection Procedure

The review process started with defining research questions as stated above. Different data sources including ACM Digital Library, IEEE explore, Springer LNCS, Google Scholar etc have been covered to gain broader knowledge. The study selection procedure is shown in Figure 1. A large number of research papers were found using different keywords (stage 1). On the basis of titles and abstracts, irrelevant or duplicate papers were excluded in stage 2 and 3 respectively. In the final stage, full text analysis was conducted and finally 61 papers were selected.

The study procedure involved a thorough literature survey of what actually is the concept of Utility Computing and the need for resource allocation and scheduling. The next step included finding out various resource management systems that are engaged in providing proficient allocation of limited resources. This was followed by an extensive survey of existing utility driven resource scheduling techniques. An investigation into the computing environment of these scheduling