International Journal of Grid and High Performance Computing
Volume 8 • Issue 1 • January-March 2016

Architectural Pattern for Scheduling Multi-Component Applications in Distributed Systems

Absalom El-Shamir Ezugwu, Federal University Lafia, Lafia, Nigeria
Marc Eduard Frincu, Faculty of Mathematics and Computer Science, West University of Timisoara, Timisoara, Romania
Sahalu Balarabe Junaidu, Ahmadu Bello University, Zaria, Nigeria

ABSTRACT

This paper presents a conceptual perspective on scheduling systems' design pattern for several classes of multi-component applications. The authors consider this scheduling problem in a wide-area network of heterogeneous computing environment. The heterogeneity in both the user application and distributed resource environments make this a challenging problem. In addition, the authors propose a component-based reference architectural model, which describes the design of a general purpose scheduling system targeted at the scheduling of multi-component applications. The design goal is to identify and map out the necessary ingredients required to effectively perform the scheduling of multi-component applications.

KEYWORDS

Multi-Agent System, Multi-Component Applications, Multi-Component Resources, Multi-Component Scheduler, Scheduling Patterns, Single-Component Scheduler

1. INTRODUCTION

The rapid innovation in distributed multi-component computing application frameworks (Hindman *et al.*, 2011; McRae, 1997), calls for an urgent need to build an equivalent multi-component distributed system infrastructure or meta-computing infrastructure. However, a number of research groups (Ghodsi *et al.,* 2011; Isard *et al.,* 2009; Grimshaw and Wulf, 1997; Foster and Kesselman, 1997) have proposed and implemented metacomputing infrastructure targeted at achieving high performance throughput for a large number of diverse compute intensive metacomputing applications. The recent shift in paradigm from parallel application, requesting for resources from single computing clusters to metacomputing applications, requesting for resources from heterogeneous metacomputing clusters can be attributed to the single goal of achieving high performance.

In this paper, we have grouped computing application into two classes based on their resource requirement needs. First is the single-component application, this is a class of application which resource requirements can be handled by resources from a single cluster. Second class of application is the multi-component application, these are sets of applications which resource requirement needs goes beyond resources provisioning from a single cluster; rather, it requires heterogeneous resources

DOI: 10.4018/IJGHPC.2016010101

from multiple clusters. These resources can include; remote databases servers, remote laboratory instruments, remote compute intensive servers, and remote network servers (Weissman, 2000). The challenges inherent in distributed heterogeneous computing environment are well known (Freund and Siegel, 1993). Examples of frameworks that provide heterogeneous computing environments include, multi-component clusters, grid and cloud computing systems (Foster and Kesselman, 1997; Weissman, 2000).

Exploiting the performance potential that comes with the heterogeneous computing environments, requires effective application scheduling. This in essence, would require the appropriate and efficient selection and allocation of candidate resources to user application. This problem is particularly challenging due to the heterogeneous and unpredictable nature of both the resources and the application itself. The problem of scheduling heterogeneous application and resource can be made more effective, by applying some scheduling heuristics that best understand the complete structures of both the application and resource information. The scheduling heuristics should be able to automatically extract this information and forward it to the global scheduler for adequate scheduling decision making.

Our intent in this paper, is to present a conceptual design framework for a multi-component based reference scheduling architecture, capable of scheduling simultaneously both single and multi-component heterogeneous applications, across diverse multi-platforms of heterogeneous multi-clusters, with the aim of reducing application execution time, achieving optimal resource throughput and utilization. Several independent or separate scheduling implementations for single-component and multi-component cluster can be seen in (Ezugwu *et al.*, 2015; Weissman, 1998; Mechoso *et al.*, 1994).

We intend to build an object-based storage infrastructure that replaces the existing Meta Directory Service (MDS) information infrastructure, which lacks the capability of supporting abstract queries from user applications (Dabhi, and Prajapati, 2008). Existing distributed systems such as the grid, uses the syntax or schema based resource matchmakers, algorithmic schedulers, and execution monitors for scripted job sequences (Dabhi, and Prajapati, 2008). To overcome the heterogeneous and dynamic nature of distributed systems, the object-based information infrastructure plays a very important role in maintaining dynamism associated with most scheduling components.

Lastly, we also intend to demonstrate how multi-component scheduling can effectively be achieved through the use of multi-agent system. This, we illustrate by presenting a simplified architectural model that result from the marriage of the two design paradigms, the object-oriented and multi-agent design paradigm together. The autonomous and computational efficiency of the multi-agent based system was equally leveraged to develop the simplified and robust multi-component scheduling model presented in the later part of this paper.

In general, a complete scheduling system should be able to accommodate the diversities in both user applications and available resources. To the best of our knowledge, none of the aforementioned related frameworks, attempt to address heterogeneity of user applications and resources concurrently. Since the common assumption is to either consider scheduling multi-component application on a single cluster or on multi cluster resources separately. In this paper, we present the architecture for a conceptual design framework for multi-component based scheduling system. The proposed scheduling framework is composed of different scheduling modules that aim at providing solutions to diverse heterogeneous applications targeting resources inside distributed multi-component environments. The extension of the proposed framework is made easier since it is modeled from the perspective of core object-oriented design paradigm. With an illustration, we show how the framework operates and used as reference architecture, starting from a user submitting an application, to scheduler selecting and allocating the best candidate resource.

The rest of the paper is organized as follows: in section 2, an overview of related work is presented. Section 3 and section 4 provide detailed discussions on the system architecture and conceptual framework model. Section 5 present architectural description of the proposed scheduling system, while Section 6 demonstrate the integration of multi-agent system into the object-oriented design model, followed by a demonstration of scenarios showing how the proposed architecture