Reputation-based dependable scheduling of workflow applications in Peer-to-Peer Grids

Mustafizur Rahman [a,*], Rajiv Ranjan [b], Rajkumar Buyya [a]

[a] Cloud Computing and Distributed Systems (CLOUDS) Laboratory, Department of Computer Science and Software Engineering,
The University of Melbourne, Victoria 3010, Australia
[b] Service Oriented Computing (SOC) Research Group, School of Computer Science and Engineering, The University of New South Wales, Sydney, Australia

ARTICLE INFO

ABSTRACT

Grids facilitate creation of wide-area collaborative environment for sharing computing or storage resources and various applications. Inter-connecting distributed Grid sites through peer-to-peer routing and information dissemination structure (also known as Peer-to-Peer Grids) is essential to avoid the problems of scheduling efficiency bottleneck and single point of failure in the centralized or hierarchical scheduling approaches. On the other hand, uncertainty and unreliability are facts in distributed infrastructures such as Peer-to-Peer Grids, which are triggered by multiple factors including scale, dynamism, failures, and incomplete global knowledge.

In this paper, a reputation-based Grid workflow scheduling technique is proposed to counter the effect of inherent unreliability and temporal characteristics of computing resources in large scale, decentralized Peer-to-Peer Grid environments. The proposed approach builds upon structured peer-to-peer indexing and networking techniques to create a scalable wide-area overlay of Grid sites for supporting dependable scheduling of applications. The scheduling algorithm considers reliability of a Grid resource as a statistical property, which is globally computed in the decentralized Grid overlay based on dynamic feedbacks or reputation scores assigned by individual service consumers mediated via Grid resource brokers. The proposed algorithm dynamically adapts to changing resource conditions and offers significant performance gains as compared to traditional approaches in the event of unsuccessful job execution or resource failure. The results evaluated through an extensive trace driven simulation show that our scheduling technique can reduce the makespan up to 50% and successfully isolate the failure-prone resources from the system.

1. Introduction

Grid computing enables the sharing, selection, and aggregation of geographically distributed heterogeneous resources, such as computational clusters, supercomputers, storage devices, and scientific instruments. These resources are under control of different Grid sites and being utilized to solve many important scientific, engineering, and business problems.

Inter-connecting distributed Grid sites through peer-to-peer routing and information dissemination structure (also known as Peer-to-Peer Grids) is essential to avoid the problems of scheduling efficiency bottleneck and single point of failure in the centralized or hierarchical scheduling approaches. Peer-to-Peer Grid (P2PG) model offers an opportunity for every site to pool its local resources as part of a single, massive scale resource sharing abstraction. P2PG infrastructures are large, heterogeneous, complex, uncertain and distributed.

* Corresponding author. Tel.: +61 3 8344 1355; fax: +61 3 9348 1184.
  E-mail addresses: mmrahman@csse.unimelb.edu.au (M. Rahman), rajiv@unsw.edu.au (R. Ranjan), raj@csse.unimelb.edu.au (R. Buyya).

In a P2PG, both control and decision making are decentralized by nature and different system components (users, services, application components) interact together to adaptively maintain and achieve a desired system wide behaviour. Furthermore, the availability, performance, and state of resources, applications and services undergo continuous changes during the life cycle of an application. Thus uncertainty and unreliability are facts in P2PG infrastructures, which are triggered by multiple factors, including: (i) software and hardware failures as the system and application scale that lead to severe performance degradation and critical information loss; (ii) dynamism (unexpected failure) that occurs due to temporal behaviours, which should be detected and resolved at runtime to cope with changing conditions; and (iii) lack of complete global knowledge that hampers efficient decision making as regards to composition and deployment of the application elements.

The aforementioned challenges are addressed in this paper by developing a novel self-managing [1] scheduling algorithm for workflow applications that takes into account the Grid site's prior performance and behaviour for facilitating opportunistic and context-aware placement of application components. The proposed scheduling algorithm is fully dependable, as it is capable of dynamically adapting to the changes in system behaviour by taking into consideration the performance metrics of Grid sites (software and hardware capability, availability, failure). The dependability of a Grid site is quantified using a decentralized reputation model, which computes local and global reputation scores for a Grid site based on the feedbacks provided by the scheduling services that have previously submitted their applications to that site. In particular, this paper contributes the following to the state-of-the-art in the Grid scheduling paradigm:

A novel Grid scheduling algorithm that aids the Grid schedulers such as resource brokers in achieving improved performance and automation through intelligent and opportunistic placement of application elements based on context awareness and dependability.

Further, the effectiveness of this contribution is appraised through:

(i) A comprehensive simulation-driven analysis of the proposed approach based on realistic and well-known application failure models to capture the transient behaviours that prevails in existing Grid-based e-Science application execution environments;

(ii) A comparative evaluation that demonstrates the self-adaptability of the proposed approach in comparison to Grid environments where: (1) resource/application behaviours do not change (i.e. no failure occurs), therefore no self-management is required and, (2) transient conditions exist but runtime systems and application elements have no capability to self-adapt.

The remainder of this paper is organized as follows. The related work that are focused on dependable application scheduling, distributed reputation models and Grid workflow management is presented in next section. Section 3 provides a brief discussion related to key system models with respect to overlay creation, application composition, task failure and application scheduling. In Section 4, we provide the distributed reputation management technique and the algorithms related to proposed dependable scheduling approach with example. Simulation setup, performance metrics and key findings of the experiments performed are analyzed and discussed in Section 5. Finally, we conclude the paper with the direction for future work.

2. Related work

The main focus of this section is to compare the novelty of the proposed work with respect to existing approaches. We classify the related research into three main areas:

2.1. Dependable scheduling

A recent work by Kim et al. [19] that advocates Content Addressable Network [28], DHT based dynamic propagation and load-balancing in desktop Grids, suffers from performance uncertainty and unreliability due to the lack of context awareness in scheduling. A most recent proposal on reputation-driven scheduling in the context of voluntary computing environments (desktop grids) has been put forward by Sonnek et al. [31]. They consider a centralized system model, where a central server is assigned responsibility for maintaining reliability ratings that form the basis for assigning tasks to group of voluntary nodes. Such centralized models for scheduling and reputation management [2] present serious bottleneck as regards to scalability of the system and autonomy of Grid sites. Moreover, these approaches are targeted on bag of tasks type of application model, whereas our approach considers scheduling of workflow applications. Currently, Grid information services [9], on which Grid schedulers [13] depend for resource selection, do not provide information regarding how the resources have performed in the recent past (performance history) or at what level they are rated by other schedulers in the system as regards to QoS satisfaction.

2.2. Distributed reputation models

There has been considerable amount of research work done in peer-to-peer (P2P) reputation systems to evaluate the trustworthiness of participating peers. These reputation systems are targeted towards P2P file sharing networks that focus on sharing and distribution of information in Internet-based environments. The PowerTrust model proposed by Zhou [39], utilizes single dimensional Overlay Hashing Functions (OHFs) for: (i) aassigning score managers for peers in the system and (ii) aggregating/computing the global reputation score. These kinds of OHFs are adequate if the search for peers/resources is based on single keyword (such as file name) or where there is single ordering in search values. However, OHFs are unable to support (or support with massive overhead) searches containing multiple keywords, range queries (such as search for a Grid site that has: Linux operating system, 100