Neurocomputing ■ (■■■■) ■■■–■■■

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Effective hybrid load scheduling of online and offline clusters for e-health service

Kai Cui [a,b], Jie Wang [a,b], Houbing Song [c], Chi Lin [a,b,*], Kuanjiu Zhou [a,b], Mingchu Li [a,b]

[a] School of Software, Dalian University of Technology, Dalian, China
[b] Key Laboratory for Ubiquitous Network and Service Software of Liaoning Province, Dalian, China
[c] West Virginia University, Virginia, United States

ARTICLE INFO

Article history:
Received 24 December 2015
Received in revised form
7 March 2016
Accepted 21 March 2016

MSC:
00-01
99-00

Keywords:
Hybrid load
Online e-health service
Compression index
Storm platform

ABSTRACT

Hybrid load in e-health services is composed of online e-health service applications and offline jobs. Previous methods overlooked the impact of system performance for the fine-grained service components. In this paper, a hybrid load scheduling scheme is proposed in which scheduling is performed not only at the level of the component, but also within components. To improve both execution efficiency and searching accuracy, the proposed algorithm searches the compressing method of the Lucene index and then filters that index. Simulations are conducted on a Storm platform to evaluate the performance of the proposed scheme. Simulation results demonstrate that the proposed scheme can increase the response speed by 67.79% with an accuracy of 95.94%, and the response speed decreases by 11.6–53.2%.

1. Instruction

Storing and managing a large scale of data and do so effectively poses a great challenge on cloud scheduling. As a promising approach for better data maintenance, the cluster system is applied in a widespread manner for the cloud services for business, medicine, and automobiles, which all generate large amounts of data. In general, a cluster system can be categorized as two types, the online current service and offline batch processing. The online service usually requires a quick response within milliseconds, while offline batch processing focuses on processing huge data in an offline manner, i.e., using MapReduce [1], diagram calculation and flow calculation, which do not require a short response time. To process the personal e-health information, the EU HealthGrid [2] achieved a variety of medical image information transmission and physiological status and other aspects of information processing. The personal health data requires to be analyzed and processed efficiently, which a typical data intensive problem [3], therefore, needs urgently to design an effective scheduling model for solving this problem [4,5].

To process the large scales of such data, the algorithm based on the hybrid Load of Online and Offline service was proposed. Srinivas et al. [6] proposed a dynamic load balancing information scheduling method with a global perspective to optimize resource utilization and task scheduling. Youngjoo et al. [7] presented the information scheduling method based on mobile-to-grid by using mobile devices and a grid method to optimize the service performance of the mobile device grid. To improve the utilization of resources and online services in the cloud data center, tasks are usually deployed via mixed batch jobs on the same physical node for clusters of mixed workloads [8]. Collins et al. [9] described the MOST method and SMART methods for building and evaluating e-health interventions, where there are approaches for systematically and efficiently optimizing behavioral interventions. Zhang et al. [10] analyzed real-time performance of the cluster, and the CPI impact on performance of CPU, which interfered to the schedule of information on the workload, and then affect the shared resources. Ahn et al. [11] proposed the method of virtual machine migration technology which accesses the resources of the hardware. The mode for sharing memory has been presented to reduce possible conflicts and then analyzes the effect from the performance of the online service. Govindan et al. [12] analyzed the performance degradation and the interference are respectively due to the workload consolidation and the shared processor

* Corresponding author at: School of Software, Dalian University of Technology, Dalian, China.
E-mail address: c.lin@dlut.edu.cn (C. Lin).

http://dx.doi.org/10.1016/j.neucom.2016.03.103
0925-2312/

caches, they presented a practical technique which works on current processor architectures and requires minimal software changes. Han et al. [13] proposed the component scheduling approach that applies an analytic performance model to predict the latencies of all components of an interactive service, and predict the performance in the scheduling method. Our scheduling approach is based on this scheme, and we optimize the method to schedule the eHealth information. In addition, literature [14] proposed the heterogeneous load collaborative filtering algorithm to execute performance prediction, and scheduled applications of service classes to clusters with smaller load interference based on prediction results. To consider the fine-grained online service components and reduce the delay, the online service overall execution performance is to being improved from the method of the literature [15].

In this paper, the approach has some similarities with the article, but it is different from the levels of scheduling is. Moreover, the literatures in the paper are considered the entire online service application to conduct the coarse-grained application scheduling for the perspective of the entire application. It is not considered the impact of the performance differences in the aspect of the fine-grained service components for the whole performance of the application.

Although, large scale data can be effectively processed, the aforementioned methods still suffer from the following problems:

(1) To process the large scale of medical information, the overall response long tail delay (LTD) is determined for the service components, which results in maximum execution time [16].
(2) In terms of system performance, processing accuracy, as well as response time analysis, methods of task scheduling still need to be improved [17,18].
(3) All approaches are designed based on scheduling and load-balancing in the experimentation, thus lacking a method to reduce scheduling task granularity [19].

A prominent drawback that also cannot be overlooked is that a large scale of the data is stored and scheduled, and in that case, the task scheduling becomes more difficult and has along response time. Therefore, enhancing the effectiveness of the scheduling in terms of hybrid load in the cluster system is of major significance for efficient e-health service.

Motivated to overcome this problem the authors develop a scheduling algorithm based on Storm platform in this paper. First, it conducts the Person E-Health Record (PEHR) index compression, and then it searches the compressed index. The objective of our scheme is to execute scheduling between and within components to guarantee the accuracy and timely response for online information processing. Moreover, this method can reduce the response LTD for fine-grained service. Finally, simulations are performed to demonstrate the advantages of the proposed scheme.

The contributions of this paper can be summarized as follows:

(1) Conduct PEHR scheduling for online and offline fine-grained services, the scheduling method to be developed among and within components in which multiple copies of requests are sent to the components for online service, while, the fastest request copy is accepted. Hence, the speed of processing the scheduling for off-line tasks are accelerated, and the delay in different rates for requesting search decreases by 11.6–53.2%, the average delay for different rates of requesting search decreases by 20–45%.
(2) With respect to response time, a request that cannot be completed within the time requirement will be re-allocated to a new component. This process makes full use of the components so as to accelerate the responding action, and internal components of the computing resource allocation algorithm under an average response time delay is decreased by 36.14–54.37%.
(3) Verify the correctness of the proposed scheme through sophisticated and extensive simulations conducted to show the performance of our scheme. Simulation results to demonstrate that the proposed scheme increases the response speed by 67.79% with an accuracy of 95.95%, and the response speed decreases by 11.6–53.2%.

The rest of the paper is organized as follows, after this introduction: in Section 2 we analyzed the scheme of the scheduling PEHR, the performance of the search component and the schedule algorithm. In Section 3, the illustrating of the performance with the scheme of the schedule algorithm is described and the Simulations are analyzed. Finally, we conclude this paper in Section 4.

2. Design and implementation

2.1. The scheduling scheme of e-health record

From the Google cluster workload trace that many longer-running jobs are seen as having relatively stable resource utilization. Thus, this research designs a component-level algorithm for e-health record, based on real-time monitoring of the performance of cluster to predict the interference of a Storm search service component [20,21]. It is proposed that it may send a redundant request to the online service application for reducing delay. Currently, there are two possible options specifically:

(1) Each request multiple copies were sent to the components of the online service and the components finally accepted the transfer request in response to the fastest results [22,23].
(2) The only request is not completed within a certain period of time. It will be resent to a new component to execute, and eventually accepted the results of the new [24,25].

More programs at low pressure service request can significantly improve the implementation performance of the service, but in the service request stressful or machine is a strong degree of interference, the above scenario would seriously reduce the performance of the online service execution [26]. The service component that gets greater interference than the next moment is assigned to relatively idle machines. This process will speed up component execution in a fine-grained way and obtain more accurate search results. The overall scheduling scheme of the e-health record is shown in Fig. 1.

2.2. Component-level algorithm of e-health scheduling

The descriptors used for a scheduling algorithm are shown in Table 1. N is a physical node in the cluster.

The scheduling algorithm is as follows:

Step 1: Obtain monitoring e-health record of the cluster machine at every scheduling interval.

Step 2: Build the performance predicting matrix $M_{NC\cdot NN}$ according to the performance predicting model.

In the matrix, $M[i][j]$ means changing the execution time after moving $Comp_i$ to the physical node $Node_j$ of the cluster. If positive, the performance is improving after moving $Comp_i$ and the search time decreases; if negative, the searching time is increasing.

Please cite this article as: K. Cui, et al., Effective hybrid load scheduling of online and offline clusters for e-health service, Neurocomputing (2016), http://dx.doi.org/10.1016/j.neucom.2016.03.103