INTERNATIONAL JOURNAL OF SCIENTIFIC PROGRESS AND RESEARCH (IJSPR) ISSN: 2349-4689

To reduce the cost minimization problem for private CDC in hybrid clouds

Mrs Ashwini [1], Mrs Tanuja[2],
BGSIT,Computer Engineering Department,
Karnataka ,INDIA

Abstract -The economy of scale provided by cloud attracts a growing number of organizations and industrial companies to deploy their applications in cloud data centers (CDCs) and to provide services to users around the world. The uncertainty of arriving tasks makes it a big challenge for private CDC to cost-effectively schedule delay bounded tasks without exceeding their delay bounds. Unlike previous studies, this paper takes into account the cost minimization problem for private CDC in hybrid clouds, where the energy price of private CDC and execution price of public clouds both show the temporal diversity. Then, this paper proposes a temporal task scheduling algorithm (TTSA) to effectively dispatch all arriving tasks to private CDC and public clouds. In each iteration of TTSA, the cost minimization problem is modeled as a mixed integer linear program and solved by a hybrid simulated-annealing particle-swarm-optimization. The experimental results demonstrate that compared with the existing methods, the optimal or suboptimal scheduling strategy produced by TTSA can efficiently increase the throughput and reduce the cost of private CDC while meeting the delay bounds of all the tasks.

Keywords—Cloud computing, cloud data center, cost delay bounded tasks, hybrid clouds, resource provisioning, task scheduling.

I. INTRODUCTION

The economy of scale provided by cloud computing has attracted many corporations to outsource their applications to cloud data center (CDC) providers [1]–[4]. In cloudcomputing, typical Infrastructure as a Service (IaaS) providers such as Rackspace [5] provide resources to support applications delivered to users. Similar to the paper in [6], from the perspective of a typical IaaS provider, private CDC inthis paper refers to a resource-limited IaaS provider that may schedule some tasks to external public clouds if its resources

cannot guarantee the expected QoS. Besides, the considerationof security and regulation causes that some applications canbe provided by private CDC only. Private CDC aims to provide services to all arriving tasks from millions of users in the most cost-effective way while ensuring user-defined delay bounds. The arrival of users' tasks is a periodic and uncertain, and therefore it is challenging for private CDC to accurately predict the upcoming tasks. Besides, the limitation of resources in private CDC makes it possible that some arriving tasks must be refused to provide delay assurance of accepted tasks when the number

of arriving tasks is unexpectedly large [7], [8]. However, this reduces the throughput of private CDC, and inevitably brings large penalty to private CDC due to the refusal of tasks. The emergence of hybrid clouds enables private CDC to outsource part of its arriving tasks to public clouds when tasks unexpectedly peak.In hybrid clouds, the total cost of private CDC mainly consists of the energy cost caused by the accepted tasks executed in it,and the execution cost of tasks dispatched to public clouds.Public clouds (e.g., Amazon EC2) deliver dynamic resources to users by creating a set of virtual machines (VMs).Delay bounded tasks usually have user-defined delay bounds to satisfy. In a real-life market, the execution price of VM instances provided by public clouds varies with the delaybounds [9]. Besides, the energy price of private CDC also shows the temporal diversity [8]. Therefore, how to minimize the total cost of private CDC in hybrid clouds where the execution and energy prices show the temporal diversity becomes a challenging problem.This work investigates the cost minimization problem for private CDC in hybrid clouds. This problem is formulated and solved by the proposed temporal task scheduling algorithm(TTSA). With the consideration of the temporal diversity in price, TTSA can effectively reduce the cost of private CDC by intelligently allocating all arriving tasks to private CDC or public clouds in their delay bounds. Then, public workload in Google production cluster [10] is adopted to evaluate the proposed TTSA. Comprehensive comparisons demonstrate that it outperforms the existing task scheduling approaches in terms of throughput and cost. The major contributions of this paper are as follows. First, the proposed method can strictly guarantee the delay bound of each delay bounded task. Second,this paper formulates an architecture of hybrid clouds that can provide temporal task scheduling. This architecture enables private CDC to outsource some of its tasks to public clouds provided that the delay bound of each arriving task is strictly ensured. Third, based on this architecture, TTSA is proposed to minimize the total

cost of private CDC by intelligently dispatching all arriving tasks in hybrid clouds.

II SYSTEM MODEL

After careful analysis the system has been identified to have the following modules:

1. Cloud Computing Module.
2. Resource Management Module.
3. Hybrid Cloud Computing Module.

1.Cloud Computing Module:

Cloud computing refers to applications and services offered over the Internet. These services are offered from data centers all over the world, which collectively are referred to as the "cloud." Cloud computing is a movement away from applications needing to be installed on an individual's computer towards the applications being hosted online. Cloud resources are usually not only shared by multiple users but as well as dynamically re-allocated as per demand. This can work for allocating resources to users in different time zones.

2. Resource Monitor Module:

Dynamic resource management has become an active area of research in the Cloud Computing paradigm. Cost of resources varies significantly depending on configuration for using them. Hence efficient management of resources is of prime interest to both Cloud Providers and Cloud Users. The success of any cloud management software critically de-pends on the flexibility; scale and efficiency with which it can utilize the underlying hardware resources while pro-viding necessary performance isolation. Successful resource management solution for cloud environments, needs to provide a rich set of resource controls for better isolation, while doing initial placement and load balancing for efficient utilization of underlying resources.

3.Hybrid Cloud Computing Module:

The architecture of hybrid clouds is shown in Fig. 1. The architecture consists of private CDC and public clouds. A great number of physical clusters in the former are virtualized to provide resources (e.g., CPU, memory, network, and storage) to users. Component Monitor watches physical clusters, and sends resource information to Scheduler. Users' tasks are first enqueued into an first-come-first-served (FCFS) queue that reports queue information to Scheduler. Besides, Predictor executes prediction algorithms [22] by using historical data to obtain future task information in private CDC and public clouds. There have been existing researches that focus on workload prediction based on historical data [23]–[27]. Therefore, this paper assumes that Predictor can well predict future information including task arriving rate, expected energy price in private CDC, expected execution price of public clouds, and expected average running time of each task in each time slot.

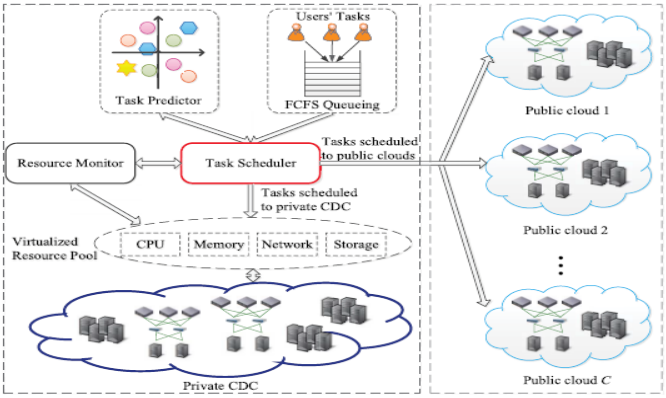

Fig. 1. Components of a hybrid cloud architecture.

Figure1 Component of a hybrid cloud architecture

III EXISTING SYSTEMS

The experimental results demonstrate that compared with the existing methods, the optimal or suboptimal scheduling strategy produced by TTSA can efficiently increase the throughput and reduce the cost of private CDC while meeting the delay bound of all the tasks.Comprehensive comparisons demonstrate that it outperforms the existing task scheduling approaches in terms of throughput and cost. Nevertheless, none of the existing studies focus on resource provisioning for delay bounded tasks in hybrid clouds. Besides, Predictor executes prediction algorithms by using historical data to obtain future task information in private CDC and public clouds. There have been existing researches that focus on workload prediction based on historical data. The temporal task scheduling problem that aims to minimize the total cost of private CDC in hybrid clouds. Similar to the work. we model private CDC as a discrete-time system evolving in a sequence of equal-length time slots. Besides, more applications are designed to execute in parallel due to the wide deployment of massive-scale commodity computers in the existing CDCs. We first clarify that the proposed TTSA algorithm focuses on tasks of parallelized applications.

IV PROPOSED SYSTEMS

This work investigates the cost minimization problem for private CDC in hybrid clouds. This problem is formulated and solved by the proposed temporal task scheduling algorithm (TTSA). With the consideration of the temporal diversity in price, TTSA can effectively reduce the cost of private CDC by intelligently

allocating all arriving tasks to private CDC or public clouds in their delay bounds.A number of methods on resource provisioning in CDCs have been proposed a lightweight system is designed to simulate real-time resource provisioning in CDCs. A virtualized system is presented to dynamically provision resources based on users' tasks. The effect of workload prediction on resource provisioning is investigated.A task scheduling method based on heuristic is proposed to maximize the profit of a private cloud while ensuring the delay bounds. Nevertheless, none of the mentioned studies considers the temporal diversity in the execution and energy prices in hybrid clouds.

V CONCLUSION

Cost minimization is an important factor for private CDC because it aims to provide services to delay bounded tasks in the most cost-effective way while guaranteeing their delay bounds. The emergence of hybrid clouds enables private CDC to meet the delay bound of each arriving task by intelligently scheduling tasks between private CDC and public clouds even if users' tasks peak unexpectedly. The temporal diversity in the execution price of public clouds, and the energy price of private CDC bring great opportunities to minimize the total cost of private CDC while guaranteeing the delay bounds of all arriving tasks. In this paper, a TTSA is proposed to dynamically schedule all arriving tasks to private CDC and public clouds intelligently. In each TTSA iteration, a mixed integer linear program is formulated to minimize the cost of private CDC in hybrid clouds. A simulated-annealing , particle-swarm-optimization combined scheduling approach is proposed to solve this program. Simulation results demonstrate that compared with some existing work, the proposed TTSA can effectively reduce the cost and improve the throughput of private CDC while guaranteeing the specified delay bounds. In the future, we would like to incorporate dispatching time and execution delay that are also important in hybrid clouds into our model. Besides, we would like to further implement realistic hybrid clouds and to evaluate the performance of the proposed TTSA. The impact of the failure/repair behaviors of clouds on the proposed approach should be explored. Some recently proposed intelligent optimization methods may be used to solve the proposed scheduling problems.

VI FUTURE SCOPES

The emergence of hybrid clouds enables private CDC to outsource part of its arriving tasks to public clouds when tasks unexpectedly peak. In hybrid clouds, the total cost of private CDC mainly consists of the energy cost caused by the accepted tasks executed in it,and the execution cost of tasks dispatched to public clouds. Delay bounded tasks usually have user-defined delay bounds to satisfy. In a real-life market, the execution price of VM instances provided by public clouds varies with the delaybounds . Besides, the energy price of private CDC also shows the temporal diversity . Therefore, how to minimize the total cost of private CDC in hybrid clouds where the execution and energy prices show the temporal diversity becomes a challenging problem.This work investigates the cost minimization problem for private CDC in hybrid clouds.

REFERENCES

[1] A. Beloglazov, J. Abawajy, and R. Buyya, "Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing," Future Gener. Comput. Syst., vol. 28, no. 5, pp. 755–768,May 2012.

[2] K. M. Sim, "Complex and concurrent negotiations for multiple interrelated e-markets," IEEE Trans. Cybern., vol. 43, no. 1, pp. 230–245,Feb. 2013.

[3] O. S. Gedik and A. A. Alatan, "3-D rigid body tracking using vision and depth sensors," IEEE Trans. Cybern., vol. 43, no. 5, pp. 1395–1405,Oct. 2013.

[4] O. Lopes, M. Reyes, S. Escalera, and J. Gonzàlez, "Spherical blurred shape model for 3-D object and pose recognition: Quantitative analysis and HCI applications in smart environments," IEEE Trans. Cybern.,vol. 44, no. 12, pp. 2379–2390, Dec. 2014.

[5] Z. Ou et al., "Is the same instance type created equal? Exploiting heterogeneity of public clouds," IEEE Trans. Cloud Comput., vol. 1, no. 2,pp. 201–214, Jul./Dec. 2013.

[6] X. Zuo, G. Zhang, and W. Tan, "Self-adaptive learning PSO-based deadline constrained task scheduling for hybrid IaaS cloud," IEEE Trans. Autom. Sci. Eng., vol. 11, no. 2, pp. 564–573, Apr. 2014.

[7] L. Wu, S. K. Garg, and R. Buyya, "SLA-based admission control for a software-as-a-service provider in cloud computing environments,"J. Comput. Syst. Sci., vol. 78, no. 5, pp. 1280–1299, Sep. 2012.

[8] J. Luo, L. Rao, and X. Liu, "Temporal load balancing with service delay guarantees for data center energy cost optimization," IEEE Trans.Parallel Distrib. Syst., vol. 25, no. 3, pp. 775–784, Mar. 2014.

[9] H. Xu and B. Li, "Dynamic cloud pricing for revenue maximization,"IEEE Trans. Cloud Comput., vol. 1, no. 2, pp. 158–171, Jul./Dec. 2013.

[10] Q. Zhang, M. F. Zhani, R. Boutaba, and J. L. Hellerstein, "Dynamic heterogeneity-aware resource provisioning in the cloud," IEEE Trans.Cloud Comput., vol. 2, no. 1, pp. 14–28, Jan./Mar. 2014.

[11] W. Tian, Y. Zhao, M. Xu, Y. Zhong, and X. Sun, "A toolkit for modeling center," IEEETrans. Autom. Sci. Eng., vol. 12, no. 1, pp. 153–161,Jan. 2015.

[12] Z. Xiao, W. Song, and Q. Chen, "Dynamic resource allocation using virtual machines for cloud computing environment," IEEE Trans. ParallelDistrib. Syst., vol. 24, no. 6, pp. 1107–1117, Jun. 2013.

[13] T. Lu, M. Chen, and L. Andrew, “Simple and effective dynamic provisioning for power-proportional data centers,” IEEE Trans. Parallel Distrib. Syst., vol. 24, no. 6, pp. 1161–1171, Jun. 2013.

[14] O. Beaumont, L. Eyraud-Dubois, C. Thraves Caro, and H. Rejeb,“Heterogeneous resource allocation under degree constraints,” IEEE Trans. Parallel Distrib. Syst., vol. 24, no. 5, pp. 926–937, May 2013.

[15] H. M. Fard, R. Prodan, and T. Fahringer, “A truthful dynamic workflow scheduling mechanism for commercial multicloud environments,” IEEETrans. Parallel Distrib. Syst., vol. 24, no. 6, pp. 1203–1212, Jun. 2013.

[16] R. N. Calheiros and R. Buyya, “Meeting deadlines of scientific workflows in public clouds with tasks replication,” IEEE Trans. Parallel Distrib. Syst., vol. 25, no. 7, pp. 1787–1796, Jul. 2014.

[17] P. Agrawal and S. Rao, “Energy-aware scheduling of distributed systems,”IEEE Trans. Autom. Sci. Eng., vol. 11, no. 4, pp. 1163–1175,Oct. 2014.

[18] Y. Yao, L. Huang, A. Sharma, L. Golubchik, and M. Neely, “Data centers power reduction: A two time scale approach for delay tolerant workloads,” in Proc. IEEE INFOCOM, Orlando, FL, USA, 2012,pp. 1431–1439.

[19] J. Cao, K. Hwang, K. Li, and A. Y. Zomaya, “Optimal multiserver configuration for profit maximization in intensive learning of Bayesian networks,” IEEE Trans.Cybern., vol. 45, no. 12, pp. 2890–2904, Dec. 2015.

cloud computing,” IEEE Trans.Parallel Distrib. Syst., vol. 24, no. 6, pp. 1087–1096, Jun. 2013.

[20] J. Bi, H. Yuan, M. Tie, and W. Tan, “SLA-based optimisation of virtualized resource for multi-tier Web applications in cloud data centres,”Enterprise Inf. Syst., vol. 9, no. 7, pp. 743–767, Nov. 2015.

[21] G. von Laszewski, J. Diaz, F. Wang, and G. C. Fox, “Comparison of multiple cloud frameworks,” in Proc. IEEE 5th Int. Conf. Cloud Comput.,Honolulu, HI, USA, 2012, pp. 734–741.

[22] E. Ramasso, M. Rombaut, and N. Zerhouni, “Joint prediction of continuous and discrete states in time-series based on belief functions,” IEEE Trans. Cybern., vol. 43, no. 1, pp. 37–50, Feb. 2013.

[23] Y. Bao, T. Xiong, and Z. Hu, “PSO-MISMO modeling strategy for multistep-ahead time series prediction,” IEEE Trans. Cybern., vol. 44,no. 5, pp. 655–668, May 2014.

[24] C.-H. Lee, F.-Y. Chang, and C.-M. Lin, “An efficient interval type-2 fuzzy CMAC for chaos time-series prediction and synchronization,”IEEE Trans. Cybern., vol. 44, no. 3, pp. 329–341, Mar. 2014.

[25] K. Yue, Q. Fang, X. Wang, J. Li, and W. Liu, “A parallel and incremental approach for data-