Transparency service model for data security in cloud computing

ABSTRACT

- Map Reduce is a software framework for processing data-intensive applications with a parallel manner in cloud computing systems.
- Some Map Reduce jobs have the deadline requirements for their job execution.
- The existing deadline-constrained Map Reduce scheduling schemes do not consider the following two problems: various node performance and dynamical task execution time.
- In this paper, we utilize the Bipartite Graph modelling to propose a new Map Reduce Scheduler called the BGMRS.

CONTINUE

- The BGMRS can obtain the optimal solution of the deadline-constrained scheduling problem by transforming the problem into a well-known graph problem: minimum weighted bipartite matching.
- The BGMRS has the following features. It considers the heterogeneous cloud computing environment, such that the computing resources of some nodes cannot meet the deadlines of some jobs.
- In addition to meeting the deadline requirement, the BGMRS also takes the data locality into the computing resource allocation for shortening the data access time of a job.

CONTINUE

- However, if the total available computing resources of the system cannot satisfy the deadline requirements of all jobs, the BGMRS can minimize the number of jobs with the deadline violation.

EXISTING SYSTEM

- Map Reduce is a software framework for processing data-intensive applications with a parallel manner in cloud computing systems.
- Some Map Reduce jobs have the deadline requirements for their job execution.
- The existing deadline-constrained Map Reduce scheduling schemes do not consider the following two problems: various node performance and dynamical task execution time.

CONTINUE

- Existing Map Reduce scheduling schemes based on multiple aspects: the quality requirements, the entities to be scheduled, and the environmental characteristics.

PROPOSED SYSTEM

- In this paper, we propose a new scheduler that utilizes the Bipartite Graph modelling to integrate the above four concerned points in the Map Reduce Scheduling.
- The proposed Map Reduce scheduler is called the BGMRS. Compared to the previous schemes, the BGMRS can dynamically adjust the map and reduce task deadlines of a job according to the execution time of already run map and reduce tasks .
- The proposed scheme will use the new map deadline to find appropriate slots for pending map tasks.

CONTINUE

- After completing all map rounds in the map phase, the final map task completion time is used to adjust the reduce deadline for all reduce tasks.
- We propose a graph-based MapReduce scheduling scheme which considers the cloud computing system with various node performance and the running jobs with different deadline requirements.
- We present an adaptive deadline partition method to dynamically regulate the execution time of the map (reduce) tasks of the job. If a map (reduce) task takes longer execution time over its task deadline, the deadlines of the pending map (reduce) tasks will be shortened to avoid violating the whole job deadline.

CONTINUE

- We transform the deadline-constrained MapReduce scheduling (DCMRS) problem as the well-known graph problem: minimum weighted bipartite matching (MWBM). Then, we proposed an efficient heuristic algorithm to solve the DCMRS problem.
- Our scheduling scheme also resolves the resource contention problem. If there are not enough resources to meet the deadlines of all concurrently running jobs, our scheme can minimize the number of deadline-over jobs.

HARDWARE REQUIREMENTS

- Processor - Pentium –III
- Speed - 1.1 Ghz
- RAM - 256 MB(min)
- Hard Disk - 20 GB
- Floppy Drive - 1.44 MB
- Key Board - Standard Windows Keyboard
- Mouse - Two or Three Button Mouse
- Monitor - SVGA

SOFTWARE REQUIREMENTS

- Operating System : Windows 8
- Front End : Java /DOTNET
- Database : Mysql/HEIDISQL

CONCLUSION

- We have investigated the Deadline-Constrained Map Reduce Scheduling (DCMRS) problem in heterogeneous cloud computing systems.
- Considering the slots in different nodes have different amount of computing resources, we first divide a job deadline into two sub-deadlines: map and reduce deadlines.
- The sub-deadlines are used for finding appropriate slots to run the tasks of the job.
- Then, we transform the DCMRS problem to a well-known Minimum Weighted Bipartite Matching (MWBM) problem.

CONTINUE

- To solve this problem, we formulate an ILP model for obtaining the optimal solution.
- We also present a heuristic algorithm involving the node group technique to decrease the computational time.

REFERENCE

[1] J. Dean and S. Ghemawat, "MapReduce: Simplified Data Processing on Large Clusters," *ACM Commun.*, vol. 51, no. 1, pp. 107–113, Jan. 2008.

[2] T. White, *Hadoop: The Definitive Guide*, 3rd ed. Inc. O'Reilly Media, 2012.

[3] X. Dong, Y. Wang, and H. Liao, "Scheduling Mixed Real-Time and Non-real-Time Applications in MapReduce Environment," in *Proc. IEEE ICPADS*, Dec. 2011, pp. 9–16.

[4] E. Hwang and K. H. Kim, "Minimizing Cost of Virtual Machines for Deadline-Constrained MapReduce Applications in the Cloud," in *Proc. ACM/IEEE GRID*, Sep. 2012, pp. 130–138.

CONTINUE

[5] J. Polo, Y. Becerra, D. Carrera, M. Steinder, I. Whalley, J. Torres, and E. Ayguade, "Deadline-Based MapReduce Workload Manage- ´ ment," *IEEE Trans. Netw. Service Manag.*, vol. 10, no. 2, pp. 231–244, 2013.

[6] B. Cho, M. Rahman, T. Chajed, I. Gupta, C. Abad, N. Roberts, and P. Lin, "Natjam: Design and Evaluation of Eviction Policies for Supporting Priorities and Deadlines in Mapreduce Clusters," in *Proc. ACM SoCC*, 2013, pp. 1–17 .

[7] Q. Chen, C. Liu, and Z. Xiao, "Improving MapReduce Performance Using Smart Speculative Execution Strategy," *IEEE Trans. Comput.*, vol. 36, no. 4, 2014.