Starving for Hints!: Toward Predicting Resource Usage Levels of Tasks in a Distributed Computing Framework Layer

Masato Asahara
NEC Green Platforms Research Laboratories

Takeshi Yoshimura
Keio University

1 Background

Massive computing power is required for data engineering and data analysis. While Hadoop brings us near-at-hand computing clusters, we face challenging issues on resource utilization efficiency for clusters.

There are many challenges to overcome with respect to the resource waste with production clusters, and it is also still an open question as to how to improve the efficiency of resource utilization with data analytics workloads that involve multiple data processings among many kind of applications. A data analytics business workload includes many relational database processings and matrix calculations, such as machine learning algorithms. The workloads require several kinds of applications and do not usually have periodicity in load increase, but most existing approaches such as [5, 4, 6] have strong assumptions regarding types of applications (e.g., web-based or heavily-data-loading applications) or workloads (e.g., periodicity in load increase). To avoid the difficulty of predicting task behaviors, some existing prediction techniques, such as Quasar [3], focus on predicting "job" behaviors, but we would argue that they potentially over- or under-collocate tasks due to differences among individual task behaviors, such as skew. Because of such differences, resource usage for individual tasks differs even when they are part of the same job.

2 Proposal

In this poster presentation, we explore the possibility of task-based resource usage prediction for efficient task scheduling. Our prediction technique has two important design features; prediction in a computing framework layer and estimation of resource usage "levels" in tasks. Computing framework layers such as MapReduce [2] contain abstract information about the behavior of running tasks [1], and it is possible to obtain from a computing framework a job name, a class name for the program, the amount of input data, and program parameters. This information enables us to extract both quantitative and qualitative characteristics of tasks even without a deep knowledge of the application code. For example, we can utilize the number of input records for a reduce-task in a MapReduce job in order to inform a task scheduler of how much memory space the task is expected to be consumed in loading the records into the memory. Information observed in the computing framework layer is used to generate a classifier for resource usage levels by means of executing a naive Bayes classification algorithm.

Another key concept is the prediction of a "level" estimation for resource usage in tasks. A level of resource usage is an indication of the degree of resources the task will require, e.g., the number of digits of memory usage. By utilizing the predicted level of resource usage, a task scheduler can assign a task on a server with an approximate necessary-amount of resources. The predicted level is also a good guide for avoiding over- or under-collocating tasks.

3 Preliminary Experiments

We have conducted preliminary experiments of our prediction technique with 692 MapReduce jobs that consist of Hive jobs issuing TPC-H queries and typical Mahout jobs, and results suggest that it offers the possibility of helping a task scheduler improve resource efficiency for a cluster and to reduce performance degradation in tasks. Prediction accuracy for CPU and heap memory usage rise as high as 81% and 70%, respectively. The accuracy of estimating the size of a file output to a local file system and to a Hadoop Distributed File System (HDFS) rise as high as 77% and 87%, respectively. The fraction of cases of performance degradation caused by wrong prediction are 11% for CPU usage, 7% for heap memory utilization, 13% for the output file size on a local file system, and 5% for that on HDFS.

References

[1] CHEN, Y., ALSPAUGH, S., AND KATZ, R. Interactive Analytical Processing in Big Data Systems: A Cross-industry Study of MapReduce Workloads. *Proc. VLDB Endow. 5*, 12 (2012), 1802–1813.

[2] DEAN, J., AND GHEMAWAT, S. MapReduce: Simplified Data Processing on Large Clusters. In *Proc. of USENIX Symp. on Opearting Systems Design and Implementation* (2004), pp. 137–150.

[3] DELIMITROU, C., AND KOZYRAKIS, C. Quasar: Resource-efficient and QoS-aware Cluster Management. In *Proc. of Int'l Conf. on Architectural Support for Programming Languages and Operating Systems* (2014), pp. 127–144.

[4] MENACHE, I., SHAMIR, O., AND JAIN, N. On-demand, Spot, or Both: Dynamic Resource Allocation for Executing Batch Jobs in the Cloud. In *Proc. of USENIX Int'l Conf. on Autonomic Computing* (2014), pp. 177–187.

[5] NARAYANAN, D., AND SATYANARAYANAN, M. Predictive Resource Management for Wearable Computing. In *Proc. of Int'l Conf. on Mobile Systems, Applications and Services* (2003), pp. 113–128.

[6] SERACINI, F., MENARINI, M., KRUEGER, I., BARESI, L., GUINEA, S., AND QUATTROCCHI, G. A Comprehensive Resource Management Solution for Web-based Systems. In *Proc. of USENIX Int'l Conf. on Autonomic Computing* (2014), pp. 233–239.