Invited Talk
Performance Modeling in MapReduce Environments: Challenges and Opportunities

Ludmila Cherkasova
Hewlett-Packard Laboratories
1501 Page Mill Road, Palo Alto, CA 94303, USA
lucy.cherkasova@hp.com

ABSTRACT

Unstructured data is the largest and fastest growing portion of most enterprise's assets, often representing 70% to 80% of online data. These steep increase in volume of information being produced often exceeds the capabilities of existing commercial databases. MapReduce and its open-source implementation Hadoop represent an economically compelling alternative that offers an efficient distributed computing platform for handling large volumes of data and mining petabytes of unstructured information. It is increasingly being used across the enterprise for advanced data analytics, business intelligence, and enabling new applications associated with data retention, regulatory compliance, e-discovery and litigation issues.

However, setting up a dedicated Hadoop cluster requires a significant capital expenditure that can be difficult to justify. Cloud computing offers a compelling alternative and allows users to rent resources in a "pay-as-you-go" fashion. For example, a list of offered Amazon Web Services includes MapReduce environment for rent. It is an attractive and cost-efficient option for many users because acquiring and maintaining complex, large-scale infrastructures is a difficult and expensive decision. One of the open questions in such environments is the amount of resources that a user should lease from the service provider. Currently, there is no available methodology to easily answer this question, and the task of estimating required resources to meet application performance goals is the solely user's responsibility. The users need to perform adequate application testing, performance evaluation, capacity planning estimation, and then request appropriate amount of resources from the service provider. To address these problems we need to understand: "What do we need to know about a MapReduce job for building an efficient and accurate modeling framework? Can we extract a representative *job profile* that reflects a set of critical performance characteristics of the underlying application during all job execution phases, i.e., map, shuffle, sort and reduce phases? What metrics should be included in the job profile?" We discuss a profiling technique for MapReduce applications that aims to construct a compact job profile that is comprised of performance *invariants* which are independent of the amount of resources assigned to the job (i.e., the size of the Hadoop cluster) and the size of the input dataset. The challenge is how to accurately predict application performance in the large production environment and for processing large datasets from

the application executions that being run in the smaller staging environment and that process smaller input datasets.

One of the major Hadoop benefits is its ability of dealing with failures (disk, processes, node failures) and allowing the user job to complete. The performance implications of failures depend on their types, when do they happen, and whether a system can offer some spare resources instead of failed ones to the running jobs. We discuss how to enhance the MapReduce performance model for evaluating the failure impact on job completion time and predicting a potential performance degradation.

Sharing a MapReduce cluster among multiple applications is a common practice in such environments. However, a key challenge in these shared environments is the ability to tailor and control resource allocations to different applications for achieving their performance goals and service level objectives (SLOs). Currently, there is no job scheduler for MapReduce environments that given a job completion deadline, could allocate the appropriate amount of resources to the job so that it meets the required SLO. In MapReduce environments, many production jobs are run periodically on new data. For example, Facebook, Yahoo!, and eBay process terabytes of data and event logs per day on their Hadoop clusters for spam detection, business intelligence and different types of optimization. For the production jobs that are routinely executed on the new datasets, can we build on-line job profiles that later are used for resource allocation and performance management by the job scheduler? We discuss opportunities and challenges for building the SLO-based Hadoop scheduler.

The accuracy of new performance models might depend on the resource contention, especially, the network contention in the production Hadoop cluster. Typically, service providers tend to over provision network resources to avoid undesirable side effects of network contention. At the same time, it is an interesting modeling question whether such a network contention factor can be introduced, measured, and incorporated in the MapReduce performance model. Benchmarking Hadoop, optimizing cluster parameter settings, designing job schedulers with different performance objectives, and constructing intelligent workload management in shared Hadoop clusters create an exciting list of challenges and opportunities for the performance analysis and modeling in MapReduce environments.

Categories and Subject Descriptors: C.4 [Computer System Organization] Performance of Systems, D.2.6.[Software] Programming Environments.

General Terms: Measurement, Performance, Design.

Keywords: MapReduce, resource allocation, measurements, workload profiling, capacity planning, job scheduling