УДК 004.24:004.6

PERFORMANCE ASSURANCE FOR BIG DATA AND CLOUD APPLICATIONS

B. ZIBITSKER
MS, PhD
and Honorable Doctor at BSUIR
President and CEO BEZNext,
Adjunct Associate Professor, DePaul University in Chicago

CEO BEZNext, Chicago, USA
DHTechnologies, Austin, TX, USA
E-mail: bzibitsker@beznext.com

Abstract. Reducing the cost and simplification of maintenance and performance management are driving factors of migrating Big Data and Data Warehouse applications to the Cloud. Migration to the cloud brings a lot of challenges. Typical questions include: How will my application work in the cloud? How much resources should be allocated in the cloud to achieve the same performance as we experience currently on prem? How to optimize resource usage and workloads performance? How to take into consideration the impact of virtualization, and what will be the impact of moving from VMs to Containers and from YARN to Kubernetes? In this presentation we will review several use cases illustrating how BEZNext Performance Assurance software incorporates ML, AI and Queueing Network Models to address these types of challenges."

Keywords: cloud, Big Data, queueing network models.

The following questions were presented at the IV international conference Big data and Advanced Analytics in 2017. In a complex Big Data environment applications compete for resources and affect each other performance. Selection of Machine Learning Algorithms and Machine Learning Libraries and Big Data YARN's Scheduler, Queues and Containers rules can significantly affect accuracy, performance and scalability of Big Data applications.

The following questions were presented at the IV international conference Big data and Advanced Analytics in 2018. Selection of the Machine Learning (ML) algorithms and ML Libraries affect accuracy, response time, scalability and success of implementing new Big Data applications. Unfortunately, algorithms providing high accuracy not necessarily provide good response time and scale well. Different algorithms take different training time and different efforts for operationalization. In this paper we will discuss results of collaborative efforts on benchmarking ML algorithms and libraries and review the algorithm of recommender selecting the appropriate ML algorithm and ML library for new Big Data applications, depending on relative importance of accuracy, response time, scalability and other criteria.