

Application Performance Management using Learning, Optimization, and Control

Xiaoyun Zhu
VMware, Inc.
Palo Alto, California, USA
xzhu@vmware.com

ABSTRACT

In the past decade, the IT industry has experienced a paradigm shift as computing resources became available as a utility through cloud based services. In spite of the wider adoption of cloud computing platforms, some businesses and organizations hesitate to move all their applications to the cloud due to performance concerns. Existing practices in application performance management rely heavily on white-box modeling and diagnosis approaches or on performance troubleshooting “cookbooks” to find potential bottlenecks and remediation steps. However, the scalability and adaptivity of such approaches remain severely constrained, especially in a highly-dynamic, consolidated cloud environment. For performance isolation and differentiation, most modern hypervisors offer powerful resource control primitives such as reservations, limits, and shares for individual virtual machines (VMs). Even so, with the exploding growth of virtual machine sprawl, setting these controls properly such that co-located virtualized applications get enough resources to meet their respective service level objectives (SLOs) becomes a nearly

insoluble task. These challenges present unique opportunities in leveraging the rich telemetry collected from applications and systems in the cloud, and in applying statistical learning, optimization, and control based techniques to developing model-based, automated application performance management frameworks. There has been a large body of research in this area in the last several years, but many problems remain. In this talk, I'll highlight some of the automated and data-driven performance management techniques we have developed, along with related technical challenges. I'll then discuss open research problems, in hope to attract more innovative ideas and solutions from a larger community of researchers and practitioners.

Categories and Subject Descriptors

D.2.9 [Software Engineering]: Management

Keywords: performance, learning, control, optimization

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage, and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s). Copyright is held by the author/owner(s).

ICPE'14, March 22–26, 2014, Dublin, Ireland.

ACM 978-1-4503-2733-6/14/03.

<http://dx.doi.org/10.1145/2568088.2576098>