

Virtualization and Databases: State of the Art and Research Challenges

Ashraf Aboulnaga
University of Waterloo
Waterloo, Canada
ashraf@cs.uwaterloo.ca

Cristiana Amza
University of Toronto
Toronto, Canada
amza@eecg.toronto.edu

Kenneth Salem
University of Waterloo
Waterloo, Canada
kmsalem@cs.uwaterloo.ca

ABSTRACT

There is currently a lot of interest in resource virtualization as an important technique for addressing the problems of manageability, reliability, and security in computer systems. Resource virtualization decouples the user's perception of hardware and software resources from the actual implementation of these resources. It adds a flexible and programmable layer of software between user applications (such as database systems) and the resources that they use. This layer of software maps the virtual resources perceived by the applications to real physical resources. An example of this layer of software is a virtual machine monitor, which partitions the resources of a machine (CPU, disk, memory, network, etc.) into multiple virtual machines, and independent operating systems and applications can be installed on each virtual machine. The power of resource virtualization comes from the ability to manage the mapping from virtual resources to physical resources in the virtualization layer,

and to change it as needed.

The trend towards virtualization is of interest to us in the database research community because database systems are increasingly being run in virtualized environments. This presents a major opportunity since virtualization can help in solving many important problems in the areas of database system usability, manageability, deployment, scalability, and availability. Leveraging the capabilities of virtualization to solve these problems will require some effort on the part of our community. At the same time, virtualization poses some unique research challenges that must be addressed to enable database systems to run efficiently in these virtualized environments that are becoming increasingly common. In this tutorial, we will introduce resource virtualization and how it affects database systems. We will present the opportunities that resource virtualization provides for database systems and the unique research challenges that it poses, and we will review ongoing research in this area.

ABOUT THE AUTHORS

Ashraf Aboulnaga (University of Waterloo, Canada). Ashraf Aboulnaga is an Assistant Professor in the David R. Cheriton School of Computer Science at the University of Waterloo. His research interests are in the area of database management, with a current focus on self-managing database systems, database systems in virtualized environments, XML databases, and data management on the web. Ashraf obtained his PhD from the University of Wisconsin in 2002. Prior to joining the University of Waterloo, he was a Research Staff Member in the Data Management Department at the IBM Almaden Research Center in San Jose, California, from 2002 to 2004.



Cristiana Amza (University of Toronto, Canada). Cristiana Amza is an assistant professor with the Department of Electrical and Computer Engineering at University of Toronto. Cristiana received her B.S. degree in Computer Engineering from Bucharest Polytechnic Institute in 1991, the M.S. and the Ph.D. degrees in Computer Science from Rice University in 1997 and 2003, respectively. Her research interests are in the design, implementation and evaluation of computer systems, with an emphasis on distributed systems and database systems. Her current work, as part of the Chameleon project at University of Toronto, focuses on infrastructure design for distributed systems that can automatically adapt to a changing environment and workload through self-managing, self-tuning and self-healing. The project goal is to provide ease of use, and differentiated quality-of-service for a range of Information Services, such as, e-commerce, on-line bidding and massively multiplayer games. The system mechanisms used in this project range from dynamic provisioning of database servers to dynamic resource partitioning within a server box. The system components are low-cost commodity components, such as, dual-processor servers, PC's and mobile devices.



Permission to make digital or hard copies of all or part 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. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

EDBT '08, March 25–30, 2008, Nantes, France.

Copyright 2008 ACM 978-1-59593-926-5/08/0003 ...\$5.00

Kenneth Salem (University of Waterloo, Canada). Ken Salem is an associate professor in the David R. Cheriton School of Computer Science at the University of Waterloo. He received his B.Sc. in electrical engineering from Carnegie Mellon University, and his Ph.D. in computer science from



Princeton University. Before coming to Waterloo, he spent several years as a member of the faculty at the University of Maryland, College Park, and he has spent sabbatical leaves at IBM's Almaden Research Center and at HP Laboratories in Palo Alto. His research interests span a variety of topics related to database management systems and storage systems.