

# Foundations of Automated Database Tuning

Surajit Chaudhuri  
Microsoft Research  
Redmond, WA 98052, USA  
surajitc@microsoft.com

Gerhard Weikum  
Max-Planck Institute for Informatics  
D-66123 Saarbruecken, Germany  
weikum@mpi-inf.mpg.de

## 1. The Challenge of Total Cost of Ownership

Our society is more dependent on information systems than ever before. However, managing the information systems infrastructure in a cost-effective manner is a growing challenge. The total cost of ownership (TCO) of information technology needs to be significantly reduced by minimizing people costs. In fact, mistakes in operations and administration of information systems are the single most reasons for system outage and unacceptable performance.

## 2. Role of Self-Managing Technology

One way of addressing the challenge of total cost of ownership is by making information systems more self-managing. In this tutorial, we will try to gain a deeper understanding of this challenge with respect to the database management systems component. A particularly difficult piece of the ambitious vision of making database systems self-managing is the automation of database performance tuning. In this tutorial, we will study the progress made thus far on this important problem. Specifically, we will try to identify the key principles and paradigms.

## 3. Goal of the Tutorial

A survey of the specific functionality, algorithms and architecture that have been deployed commercially to enable automated database tuning is one possible way to understand this space. A tutorial at VLDB 2004 addressed this need. However, we will be adding several unique elements that were missing in the 2004 tutorial to further enrich the scope of the tutorial at VLDB 2006.

First, we will distinguish the foundational concepts from specific implementation of the features in individual products. Lack of a broad understanding of the foundational issues reduces the exposition to a laundry list of isolated features without any underlying principles. It also makes it impossible to develop a framework to compare and contrast existing technology. Such a

Permission to copy without fee all or part of this material is granted provided that the copies are not made or distributed for direct commercial advantage, the VLDB copyright notice and the title of the publication and its date appear, and notice is given that copying is by permission of the Very Large Data Base Endowment. To copy otherwise, or to republish, to post on servers or to redistribute to lists, requires a fee and/or special permission from the publisher, ACM.

VLDB '06, September 12–15, 2006, Seoul, Korea.

Copyright 2006 VLDB Endowment, ACM 1-59593-385-9/06/09

framework can play an important role in explicating the implicit assumptions, strengths as well as weaknesses in existing technology. Indeed, identifying the weaknesses could be one of the most valuable contributions of a tutorial as it opens the door for future research.

Building such a foundation requires us to depend on a number of mathematical tools such as queuing theory, combinatorial optimization and other algorithmic techniques. We will leverage such tools to characterize each of the self-manageability challenges and their solutions.

Next, we recognize from our experience that a tutorial on self-manageability is intricately tied to the resource that is being managed. No evaluation of the self-manageability is possible without an understanding of the use of the resource (e.g., memory, access paths) in a database management system. Relating the use of the resource in a relational database management system with its manageability challenges is a key requirement for gaining insight. Such a discussion will also point to how there may be an opportunity and in fact a need to consider changing our database server implementation itself (in some cases simplifying) to reduce the total cost of ownership.

Finally, there has been substantial work in the academic world, past as well as recent, that relate to self-managing DBMS technology. Our tutorial will include thorough discussion of past academic work.

## 4. Short Bios

*Surajit Chaudhuri* leads the Data Management and Exploration Group in Microsoft Research. He started the AutoAdmin research project in 1996 focusing on self-tuning database systems. Index Tuning Wizard (SQL Server 7.0 and SQL Server 2000) and Database Tuning Advisor in SQL Server 2005 are built using technology from the AutoAdmin project.

*Gerhard Weikum* is the director of the Database and Information Systems group at the Max-Planck Institute for Informatics in Saarbruecken, Germany. He was awarded the VLDB 2002 “ten-year best paper” award for his work on automated tuning.