

Self Tuning Databases

Projects from Microsoft and IBM

Sections

1. Microsoft's Database Research Group
 1. Situation and Vision
 2. The AutoAdmin Project
 3. Strategies for Index Selection
 4. Components of AutoAdmin
 5. Results
2. IBM Autonomic Computing
 1. SMART Project
 2. Smart Components
3. Summary
4. References

19.11.2003

Seminar Self Tuning Databases
Projects from Microsoft and IBM

2

1. Microsoft's Database Research Group

■ 3 main research fields

- self tuning databases
 - AutoAdmin Project, focussing on
 - index selection
 - materialized views
- robust applications
 - Phoenix Project, focussing on
 - improving application availability
 - error handling robustness
 - database recovery techniques
- semantic caching and multi-query optimization
 - QP Recycler Project

19.11.2003

Seminar Self Tuning Databases
Projects from Microsoft and IBM

3

1.1. Situation and Vision

- databases = integral components for enterprises
- TCO important, but increasing through
 - great administration efforts
 - system tuning = complex tasks
 - much expertise needed
- „The vision is to have a system that consistently delivers high performance with little or no administration, regardless of changes in its load and environment.“
(from Microsoft Database Research Group)

19.11.2003

Seminar Self Tuning Databases
Projects from Microsoft and IBM

4

1.2. The AutoAdmin Project

- long-term project
- in short-term focussing on physical DB design
 - materialized views
 - index selection problem
 - studied since 70's, but few research prototypes or commercial products, because of
 - complex workload, changing over time
 - modern query processing systems exploit indexes in more ways than early relational systems (multi-column indexes, index-only queries ...)
 - characteristics of query optimizers → specific query plans

19.11.2003

Seminar Self Tuning Databases
Projects from Microsoft and IBM

5

1.3. Strategies for Index Selection

- 3 approaches
 - „textbook“ approach
 - semantic information and simple statistic methods used for index selection
 - but no consideration of workload
 - „expert system“ approach
 - knowledge of good design as rules for index selection
 - workload can be considered
 - but no connection to query optimizer of database
 - optimizer must use selected indexes
 - „cost estimates“ approach (Microsoft and IBM)
 - optimizer establishes costs for indexes
 - index selection tool compares costs regarding workload
 - no split between optimizer and index selection tool

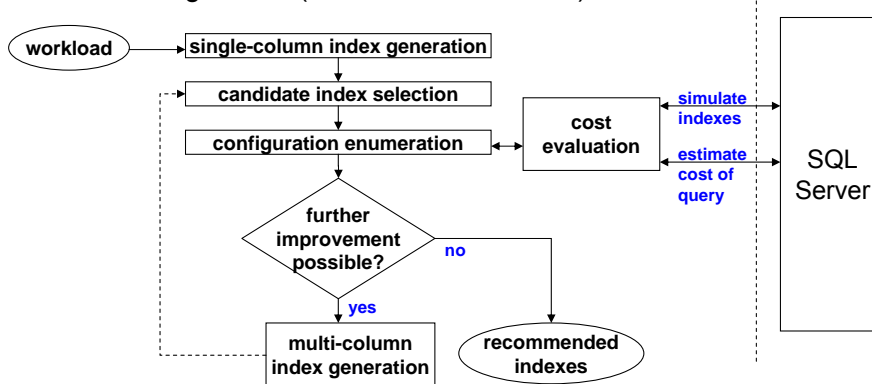
19.11.2003

Seminar Self Tuning Databases
Projects from Microsoft and IBM

6

1.4. Components of AutoAdmin

- index analysis utility (about 1998)
 - creates set of hypothetical indexes regarding workload
 - analyzes their impact
- index tuning wizard (since SQL Server 7.0)



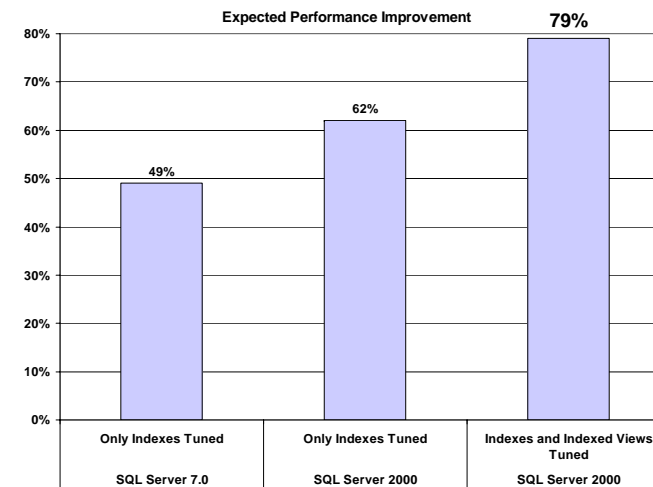
19.11.2003

Seminar Self Tuning Databases
Projects from Microsoft and IBM

7

1.5. Results of AutoAdmin

- sample workload on 1,2 GB database



Warning: Numbers provided by Microsoft

19.11.2003

Seminar Self Tuning Databases
Projects from Microsoft and IBM

8

2. IBM Autonomic Computing

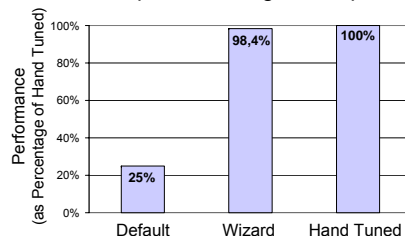
- Autonomic Computing project
 - eLiza Project = server part
 - SMART project (Self-Managing And Resource Tuning)
 - aims to make DB2 more autonomic
 - part of Autonomic Computing project
- „The continued growth of the IT industry demands that we automate many of the functions that are now performed by humans, both to save cost and to penetrate new markets.“ (from *SMART: Making DB2 (more) Autonomic*)

2.1. SMART Project Overview

- founded 2000
- 2 IBM research centers, 2 IBM development labs and over 6 universities involved
- „Our approach is to „walk before we run“, by developing tools that ease the DBA’s load, but keep the DBA in the loop. Once we earn the DBA’s trust, we can increasingly automate the function.“ (from *SMART: Making DB2 (more) Autonomic*)

2.2. SMART Components (1)

- 10 major components in DB2 UDB
 - Query Optimizer (LEarning Optimizer, LEO)
 - uses cost models
 - automatically rewrites queries
 - makes use of materialized views, etc. ...
 - Configuration Advisor
 - initial capacity planning, installation, configuration, deployment, etc.
 - sets 36 of most important configuration parameters



Warning: Numbers provided by IBM

2.2. SMART Components (2)

- Automatic Index Reorganisation
 - eliminates empty and fragmented pages
 - includes automatic index page merging mechanism (neighborhood leaf nodes merged)
- Design Advisor
 - recommends indexes for tables
 - based upon workload analysis
 - similar to Index Tuning Wizard from Microsoft
- Automatic Query Parallelism Selection
 - uses SMP architecture of server
 - much overhead (communication between 2 or more CPUs)
 - simple queries: one CPU used
 - complex queries: more CPUs used

2.2. SMART Components (3)

- Automatic Load Tuning Utility
 - makes optimal selections on memory consumption, I/O parallelism, etc. ...
- Query Patroller
 - prioritizes and schedules queries depending on load
 - guided by policies (made by DBA)
- Automatic Incremental Restore
 - allows DBA to specify an incremental image to restore
 - uses DB history to find images needed to complete restore process

2.2. SMART Components (4)

- Automatic Consistency Checking
 - protects integrity of data
 - detects corrupted data from incomplete I/O when reading from disk
- Support Serviceability Utility
 - monitoring and diagnostic functions
 - collected data stored in HTML
 - context-sensitive, interactive help desk

3. Summary

- administration effort reduced → TCOs reduced
- a lot of advances in relative short time, but just beginning
- much more components to come
- Microsoft
 - concentrating on database aspects
 - AutoAdmin Project, until now only focus on physical DB design
 - <http://research.microsoft.com/dmx/autoadmin/>
- IBM
 - Autonomic Computing: self tuning and self management components in all products
 - <http://www-3.ibm.com/autonomic/index.shtml>
 - SMART = database part of autonomic computing
 - covering „all“ facets of DB Administration
 - much more extensive research than Microsoft

4. References

- David Lomet, Roger Barga, Surajit Chaudhuri, Paul Larson
„The Microsoft Database Research Group“
- Surajit Chaudhuri, Vivek Nasarayya
„An Efficient, Cost-driven Index Selection Tool for MS SQL Server“
- Surajit Chaudhuri, Vivek Narasayya
„AutoAdmin „What-if“ Index Analysis Utility“
- Guy M. Lohman, Sam S. Lightstone
„SMART: Making DB2 (more) Autonomic“
- Sam S. Lightstone, Guy M. Lohman, Danny Zilio
„Toward Autonomic Computing with DB2 Universal Database“
- Eva Kwan, Sam S. Lightstone, Adam Storm, Leanne Wu
„Automatic Configuration for IBM DB2 Universal Database“
- WebSites (mentioned before)



The End

Any Questions?