



DATABASE ADMINISTRATION

The Complete Guide to DBA Practices and Procedures

SECOND EDITION

CRAIG S. MULLINS

Accolades for *Database Administration*

“I’ve forgotten how many times I’ve recommended this book to people. It’s well written, to the point, and covers the topics that you need to know to become an effective DBA.”

—Scott Ambler, Thought Leader, Agile Data Method

“This is a well-written, well-organized guide to the practice of database administration. Unlike other books on general database theory or relational database theory, this book focuses more directly on the theory and reality of database administration as practiced by database professionals today, and does so without catering too much to any specific product implementation. As such, *Database Administration* is very well suited to anyone interested in surveying the job of a DBA or those in similar but more specific roles such as data modeler or database performance analyst.”

—Sal Ricciardi, Program Manager, Microsoft

“One of Craig’s hallmarks is his ability to write in a clear, easy-to-read fashion. The main purpose of any technical book is to transfer information from writer to reader, and Craig has done an excellent job. He wants the reader to learn—and it shows.”

—Chris Foot, Manager, Remote DBA Experts and Oracle ACE

“A complete and comprehensive listing of tasks and responsibilities for DBAs, ranging from creating the database environment to data warehouse administration, and everything in between.”

—Mike Tarrani, Computer Consultant

“I think every business manager and every IT manager should have a copy of this book.”

—Dan Hotka, Independent Consultant and Oracle ACE

“This book by Craig Mullins is wonderfully insightful and truly important. Mullins describes the role and duties of data administrators and database administrators in modern organizations with remarkable insight and clarity.”

—Michael Tozer, Author and former U.S. Navy officer

Database Administration: The Complete Guide to DBA Practices and Procedures

Table of Contents

Contents

Preface

How to Use This Book

Acknowledgments

About the Author

Chapter 1 What Is a DBA?

Why Learn Database Administration?

A Unique Vantage Point

DBA Salaries

Database Technology

The Management Discipline of Database Administration

A Day in the Life of a DBA

Evaluating a DBA Job Offer

Database, Data, and System Administration

Data Administration

Database Administration

System Administration

DBA Tasks

Database Design

Performance Monitoring and Tuning

Ensuring Availability

Database Security and Authorization

Table of Contents

Governance and Regulatory Compliance

Backup and Recovery

Ensuring Data Integrity

DBMS Release Migration

Jack-of-All-Trades

The Types of DBAs

System DBA

Database Architect

Database Analyst

Data Modeler

Application DBA

Task-Oriented DBA

Performance Analyst

Data Warehouse Administrator

Staffing Considerations

How Many DBAs?

DBA Reporting Structures

Multiplatform DBA Issues

Production versus Test

The Impact of Newer Technology on DBA

Procedural DBAs: Managing Database Logic

The Internet: From DBA to eDBA

The Personal DBA and the Cloud

NoSQL, Big Data, and the DBA

New Technology Impacts on DBA

DBA Certification

The Rest of the Book

Review

Table of Contents

Bonus Question

Chapter 2 Creating the Database Environment

Defining the Organizations DBMS Strategy

Choosing a DBMS

DBMS Architectures

DBMS Clustering

DBMS Proliferation

Hardware Issues

Cloud Database Systems

Installing the DBMS

DBMS Installation Basics

Hardware Requirements

Storage Requirements

Memory Requirements

Configuring the DBMS

Connecting the DBMS to Supporting Infrastructure Software

Installation Verification

DBMS Environments

Upgrading DBMS Versions and Releases

Features and Complexity

Complexity of the DBMS Environment

Reputation of the DBMS Vendor

Support Policies of the DBMS

Organization Style

DBA Staff Skill Set

Platform Support

Supporting Software

Fallback Planning

Migration Verification

Table of Contents

The DBMS Upgrade Strategy

Database Standards and Procedures

Database Naming Conventions

Other Database Standards and Procedures

DBMS Education

Summary

Review

Bonus Question

Suggested Reading

Chapter 3 Data Modeling and Normalization

Data Modeling Concepts

Entity-Relationship Diagramming

The Components of a Data Model

Entities

Attributes

Keys

Relationships

Discovering Entities, Attributes, and Relationships

Conceptual, Logical, and Physical Data Models

What Is Normalization?

The Normal Forms

First Normal Form

Second Normal Form

Third Normal Form

A Normalized Data Model

Further Normal Forms

Normalization in Practice

Table of Contents

Additional Data Modeling Issues

Summary

Review

Bonus Question

Suggested Reading

Chapter 4 Database Design

From Logical Model to Physical Database

Transform Entities to Tables

Transform Attributes to Columns

Build Referential Constraints for All Relationships

Build Physical Data Structures

Database Performance Design

Designing Indexes

Hashing

Clustering

Interleaving Data

Denormalization

When to Denormalize

Prejoined Tables

Report Tables

Mirror Tables

Split Tables

Combined Tables

Redundant Data

Repeating Groups

Derivable Data

Hierarchies

Special Physical Implementation Needs

Table of Contents

Denormalization Summary

Views

Data Definition Language

Temporal Data Support

A Temporal Example

Business Time and System Time

Summary

Review

Bonus Question

Suggested Reading

Chapter 5 Application Design

Database Application Development and SQL

SQL

Set-at-a-Time Processing and Relational Closure

Embedding SQL in a Program

SQL Middleware and APIs

Application Infrastructure

Object Orientation and SQL

Types of SQL

SQL Coding for Performance

Querying XML Data

Defining Transactions

Transaction Guidelines

Unit of Work

Transaction Processing Systems

Application Servers

Locking

Types of Locks

Table of Contents

Lock Time-outs

Deadlocks

Lock Duration

Lock Escalation

Programming Techniques to Minimize Locking Problems

Locking Summary

Batch Processing

Summary

Review

Bonus Question

Suggested Reading

Chapter 6 Design Reviews

What Is a Design Review?

Rules of Engagement

Design Review Participants

Knowledge and Skills Required

Types of Design Reviews

Conceptual Design Review

Logical Design Review

Physical Design Review

Organizational Design Review

SQL and Application Code Design Review

Pre-Implementation Design Review

Post-Implementation Design Review

Design Review Output

Additional Considerations

Dealing with Remote Staff

Mentorship and Knowledge Transfer

Table of Contents

Summary

Review

Suggested Reading

Chapter 7 Database Change Management

Change Management Requirements

The Change Management Perspective of the DBA

Types of Changes

DBMS Software

Hardware Configuration

Logical and Physical Design

Applications

Physical Database Structures

Impact of Change on Database Structures

The Limitations of ALTER

Database Change Scenarios

Comparing Database Structures

Requesting Database Changes

Standardized Change Requests

Communication

Coordinating Database and Application Changes

Compliance

DBA Scripts and Change Management

Summary

Review

Suggested Reading

Chapter 8 Data Availability

Defining Availability

Table of Contents

Increased Availability Requirements

Cost of Downtime

How Much Availability Is Enough?

Availability Problems

Loss of the Data Center

Network Problems

Loss of the Server Hardware

Disk-Related Outages

Operating System Failure

DBMS Software Failure

Application Problems

Security and Authorization Problems

Corruption of Data

Loss of Database Objects

Loss of Data

Data Replication and Propagation Failures

Severe Performance Problems

Recovery Issues

DBA Mistakes

Outages: Planned and Unplanned

Ensuring Availability

Perform Routine Maintenance While Systems Remain Operational

Automate DBA Functions

Exploit High-Availability Features

Exploit Clustering Technology

Database Architecture and NoSQL

Summary

Review

Suggested Reading

Table of Contents

Chapter 9 Performance Management

Defining Performance

A Basic Database Performance Road Map

Monitoring versus Management

Reactive versus Proactive

Preproduction Performance Estimation

Historical Trending

Service-Level Management

Types of Performance Tuning

System Tuning

Database Tuning

Application Tuning

Performance Tuning Tools

DBMS Performance Basics

Summary

Review

Bonus Question

Suggested Reading

Chapter 10 System Performance

The Larger Environment

Interaction with the Operating System

Allied Agents

Hardware Configuration

Components of the DBMS

DBMS Installation and Configuration Issues

Types of Configuration

Memory Usage

Table of Contents

Data Cache Details

Open Database Objects

Database Logs

Locking and Contention

The System Catalog

Other Configuration Options

General Advice

System Monitoring

Summary

Review

Bonus Question

Suggested Reading

Chapter 11 Database Performance

Techniques for Optimizing Databases

Partitioning

Raw Partition versus File System

Indexing

Denormalization

Clustering

Interleaving Data

Free Space

Compression

File Placement and Allocation

Page Size (Block Size)

Database Reorganization

Determining When to Reorganize

Automation

Summary

Table of Contents

Review

Suggested Reading

Chapter 12 Application Performance

Designing Applications for Relational Access

Relational Optimization

CPU and I/O Costs

Database Statistics

Query Analysis

Joins

Access Path Choices

Additional Optimization Considerations

View Access

Query Rewrite

Rule-Based Optimization

Reviewing Access Paths

Forcing Access Paths

SQL Coding and Tuning for Efficiency

A Dozen SQL Rules of Thumb

Additional SQL Tuning Tips

Identifying Poorly Performing SQL

Summary

Review

Suggested Reading

Chapter 13 Data Integrity

Types of Integrity

Database Structure Integrity

Types of Structural Problems

Table of Contents

Managing Structural Problems

Semantic Data Integrity

Entity Integrity

Unique Constraints

Data Types

Default Values

Check Constraints

Triggers

Referential Integrity

Temporal Database Systems

Summary

Review

Suggested Reading

Chapter 14 Database Security

Data Breaches

Database Security Basics

Database Users

Granting and Revoking Authority

Types of Privileges

Granting to PUBLIC

Revoking Privileges

Label-Based Access Control

Security Reporting

Authorization Roles and Groups

Roles

Groups

Other Database Security Mechanisms

Using Views for Security

Table of Contents

Using Stored Procedures for Security

Encryption

Data at Rest Encryption

Data in Transit Encryption

Encryption Techniques

SQL Injection

SQL Injection Prevention

Auditing

External Security

Job Scheduling and Security

Non-DBMS DBA Security

DBMS Fixpacks and Maintenance

Summary

Review

Suggested Reading

Chapter 15 Regulatory Compliance and Database

Administration

A Collaborative Approach to Compliance

Why Should DBAs Care about Compliance?

Metadata Management, Data Quality, and Data Governance

Metadata

Data Quality

Data Governance

Database Auditing and Data Access Tracking

Database Auditing Techniques

Privileged User Auditing

Data Masking and Obfuscation

Table of Contents

Data Masking Techniques

Database Archiving for Long-Term Data Retention

The Life Cycle of Data

Database Archiving

Components of a Database Archiving Solution

The Impact of e-Discovery on DBA

Closer Tracking of Traditional DBA Tasks

Database Change Management

Database Backup and Recovery

Summary

Review

Suggested Reading

Chapter 16 Database Backup and Recovery

The Importance of Backup and Recovery

Preparing for Problems

Backup

Full versus Incremental Backups

Database Objects and Backups

DBMS Control

Concurrent Access Issues

Backup Consistency

Log Archiving and Backup

Determining Your Backup Schedule

DBMS Instance Backup

Designing the DBMS Environment for Recovery

Alternate Approaches to Database Backup

Document Your Backup Strategy

Database Object Definition Backups

Table of Contents

Recovery

- Determining Recovery Options
- General Steps for Database Object Recovery
- Types of Recovery
- Index Recovery
- Testing Your Recovery Plan
- Recovering a Dropped Database Object
- Recovering Broken Blocks and Pages
- Populating Test Databases

Alternatives to Backup and Recovery

- Standby Databases
- Replication
- Disk Mirroring

Summary

Review

Suggested Reading

Chapter 17 Disaster Planning

The Need for Planning

- Risk and Recovery

General Disaster Recovery Guidelines

- The Remote Site
- The Written Plan
- Personnel

Backing Up the Database for Disaster Recovery

- Tape Backups
- Storage Management Backups
- Other Approaches
- Some Guidelines

Table of Contents

Disaster Prevention

Disaster and Contingency Planning Web Sites

Summary

Review

Suggested Reading

Chapter 18 Data and Storage Management

Storage Management Basics

Files and Data Sets

File Placement on Disk

Raw Partitions versus File Systems

Temporary Database Files

Space Management

Data Page Layouts

Index Page Layouts

Transaction Logs

Fragmentation and Storage

Storage Options

RAID

JBOD

Storage Area Networks

Network-Attached Storage

Tiered Storage

Planning for the Future

Capacity Planning

Summary

Review

Suggested Reading

Table of Contents

Chapter 19 Data Movement and Distribution

Loading and Unloading Data

The LOAD Utility

The UNLOAD Utility

Maintaining Application Test Beds

EXPORT and IMPORT

Bulk Data Movement

ETL Software

Replication and Propagation

Messaging Software

Other Methods

Distributed Databases

Setting Up a Distributed Environment

Data Distribution Standards

Accessing Distributed Data

Two-Phase COMMIT

Distributed Performance Problems

Summary

Review

Bonus Question

Suggested Reading

Chapter 20 Data Warehouse Administration

What Is a Data Warehouse?

Analytical versus Transaction Processing

Administering the Data Warehouse

Too Much Focus on Technology?

Data Warehouse Design

Data Movement

Table of Contents

- Data Cleansing
- Data Warehouse Scalability
- Data Warehouse Performance
- Data Freshness
- Data Content
- Data Usage
- Financial Chargeback
- Backup and Recovery
- Dont Operate in a Vacuum!

Summary

Review

Suggested Reading

Chapter 21 Database Connectivity

Multitier, Distributed Computing

- A Historical Look
- Business Issues
- What Is Client/Server Computing?
- Types of Client/Server Applications

Network Traffic

- Database Gateways
- Database Drivers
- Connection Pooling

Databases, the Internet, and the Web

- Internet-Connected Databases
- Web Development and Web Services

Summary

Review

Suggested Reading

Table of Contents

Chapter 22 Metadata Management

What Is Metadata?

From Data to Knowledge and Beyond

Metadata Strategy

Data Warehousing and Metadata

Types of Metadata

Repositories and Data Dictionaries

Repository Benefits

Repository Challenges

Data Dictionaries

Summary

Review

Suggested Reading

Chapter 23 DBA Tools

Types and Benefits of DBA Tools

Data Modeling and Design

Database Change Management

Table Editors

Performance Management

Backup and Recovery

Database Utilities

Data Protection, Governance, Risk, and Compliance Tools

Data Warehousing, Analytics, and Business Intelligence

Programming and Development Tools

Miscellaneous Tools

Examine Native DBA Tools

Evaluating DBA Tool Vendors

Homegrown DBA Tools

Table of Contents

Summary

Review

Chapter 24 DBA Rules of Thumb

Write Down Everything

Keep Everything

Automate!

Share Your Knowledge

Analyze, Simplify, and Focus

Dont Panic!

Measure Twice, Cut Once

Understand the Business, Not Just the Technology

Dont Become a Hermit

Use All of the Resources at Your Disposal

Keep Up-to-Date

Invest in Yourself

Summary

Final Exam

Appendix A: Database Fundamentals

What Is a Database?

Why Use a DBMS?

Advantages of Using a DBMS

Summary

Appendix B: The DBMS Vendors

The Big Three

The Second Tier

Other Significant Players

Table of Contents

Open-Source DBMS Offerings

Nonrelational DBMS Vendors

NoSQL DBMS Vendors

Object-Oriented DBMS Vendors

PC-Based DBMS Vendors

Appendix C: DBA Tool Vendors

The Major Vendors

Other DBA Tool Vendors

Data Modeling Tool Vendors

Repository Vendors

Data Movement and Business Intelligence Vendors

Appendix D: DBA Web Resources

Usenet Newsgroups

Mailing Lists

Web Sites, Blogs, and Portals

Vendor Web Sites

Magazine Web Sites

Consultant Web Sites

Blogs

Database Portals

Other Web Sites

Appendix E: Sample DBA Job Posting

Job Posting

Database Administrator (DBA)

Bibliography

Database Management and Database Systems

Table of Contents

Data Administration, Data Modeling, and Database Design

Database Security, Protection, and Compliance

Data Warehousing

SQL

Object Orientation and Database Management

Operating Systems

Related Topics

DB2

IMS

MySQL

Oracle

SQL Server

Sybase

Other Database Systems

Glossary

A

B

C

D

E

F

G

H

I

J

K

Table of Contents

L

M

N

O

P

Q

R

S

T

U

V

W

X

Z

Index