

# Microsoft SQL Server 2012 Internals

Kalen Delaney

Published with the authorization of Microsoft Corporation by:  
O'Reilly Media, Inc.  
1005 Gravenstein Highway North  
Sebastopol, California 95472

Copyright © 2013 by Kalen Delaney  
All rights reserved. No part of the contents of this book may be reproduced or transmitted in any form or by any means without the written permission of the publisher.

ISBN: 978-0-7356-5856-1

1 2 3 4 5 6 7 8 9 LSI 8 7 6 5 4 3

Printed and bound in the United States of America.

Microsoft Press books are available through booksellers and distributors worldwide. If you need support related to this book, email Microsoft Press Book Support at [mspinput@microsoft.com](mailto:mspinput@microsoft.com). Please tell us what you think of this book at <http://www.microsoft.com/learning/booksurvey>.

Microsoft and the trademarks listed at <http://www.microsoft.com/about/legal/en/us/IntellectualProperty/Trademarks/EN-US.aspx> are trademarks of the Microsoft group of companies. All other marks are property of their respective owners.

The example companies, organizations, products, domain names, email addresses, logos, people, places, and events depicted herein are fictitious. No association with any real company, organization, product, domain name, email address, logo, person, place, or event is intended or should be inferred.

This book expresses the author's views and opinions. The information contained in this book is provided without any express, statutory, or implied warranties. Neither the authors, O'Reilly Media, Inc., Microsoft Corporation, nor its resellers, or distributors will be held liable for any damages caused or alleged to be caused either directly or indirectly by this book.

**Acquisitions and Developmental Editor:** Michael Bolinger

**Production Editor:** Kara Ebrahim

**Editorial Production:** Box 12 Communications

**Technical Reviewers:** Benjamin Nevarez and Jonathan Kehayias

**Copyeditor:** Box 12 Communications

**Indexer:** Box 12 Communications

**Cover Design:** Twist Creative • Seattle

**Cover Composition:** Karen Montgomery

**Illustrator:** Rebecca Demarest

# Contents at a glance

	<i>Introduction</i>	<i>xix</i>
CHAPTER 1	SQL Server 2012 architecture and configuration	1
CHAPTER 2	The SQLOS	35
CHAPTER 3	Databases and database files	99
CHAPTER 4	Special databases	139
CHAPTER 5	Logging and recovery	171
CHAPTER 6	Table storage	203
CHAPTER 7	Indexes: internals and management	297
CHAPTER 8	Special storage	381
CHAPTER 9	Special indexes	457
CHAPTER 10	Query execution	513
CHAPTER 11	The Query Optimizer	611
CHAPTER 12	Plan caching and recompilation	703
CHAPTER 13	Transactions and concurrency	765
CHAPTER 14	DBCC internals	837
	<i>Index</i>	<i>903</i>



# Contents

<i>Introduction</i> .....	<i>xix</i>
<b>Chapter 1 SQL Server 2012 architecture and configuration</b>	<b>1</b>
SQL Server editions .....	1
SQL Server installation and tools .....	2
SQL Server metadata .....	3
Compatibility views .....	3
Catalog views .....	4
Dynamic Management Objects .....	6
Other metadata .....	7
Components of the SQL Server engine .....	10
Protocols .....	11
Query processor .....	12
The storage engine .....	14
SQL Server 2012 configuration .....	17
Using SQL Server Configuration Manager .....	18
Managing services .....	19
SQL Server system configuration .....	21
Operating system configuration .....	21
Trace flags .....	23
SQL Server configuration settings .....	24
Conclusion .....	33

---

## What do you think of this book? We want to hear from you!

Microsoft is interested in hearing your feedback so we can continually improve our books and learning resources for you. To participate in a brief online survey, please visit:

[microsoft.com/learning/booksurvey](http://microsoft.com/learning/booksurvey)

<b>Chapter 2</b>	<b>The SQLOS</b>	<b>35</b>
	NUMA architecture . . . . .	36
	The scheduler . . . . .	37
	Understanding SQL Server schedulers . . . . .	38
	Binding schedulers to CPUs . . . . .	41
	Observing scheduler internals . . . . .	42
	Understanding the Dedicated Administrator Connection (DAC) . . . . .	45
	Memory . . . . .	47
	The buffer pool and the data cache . . . . .	47
	Column store object pool . . . . .	48
	Access to in-memory data pages . . . . .	48
	Page management in the data cache . . . . .	48
	The free buffer list and the lazywriter . . . . .	49
	Checkpoints . . . . .	50
	Memory management in other caches . . . . .	52
	The Memory Broker . . . . .	54
	Memory sizing . . . . .	54
	Buffer pool sizing . . . . .	55
	SQL Server Resource Governor . . . . .	61
	Resource Governor overview . . . . .	61
	Resource Governor controls . . . . .	70
	Resource Governor metadata . . . . .	71
	Extended Events . . . . .	73
	Extended Events architecture . . . . .	73
	Event execution life cycle . . . . .	73
	Core concepts . . . . .	75
	Extended Events DDL and querying . . . . .	83
	Extended Events UI . . . . .	86
	Conclusion . . . . .	97
<b>Chapter 3</b>	<b>Databases and database files</b>	<b>99</b>
	Working with sample databases . . . . .	100
	AdventureWorks . . . . .	100
	pubs . . . . .	101

Northwind . . . . .	101
Understanding database files . . . . .	101
Creating a database . . . . .	104
Using CREATE DATABASE: an example . . . . .	106
Expanding or shrinking a database . . . . .	106
Automatic file expansion . . . . .	106
Manual file expansion . . . . .	107
Fast file initialization . . . . .	107
Automatic shrinkage . . . . .	108
Manual shrinkage . . . . .	108
Using database filegroups . . . . .	109
The default filegroup . . . . .	110
A FILEGROUP CREATION example . . . . .	112
Filestream filegroups . . . . .	113
Altering a database . . . . .	114
ALTER DATABASE examples . . . . .	115
Databases under the hood . . . . .	116
Space allocation . . . . .	116
Setting database options . . . . .	119
State options . . . . .	122
Cursor options . . . . .	125
Auto options . . . . .	125
SQL options . . . . .	126
Database recovery options . . . . .	128
Other database options . . . . .	129
Understanding database security . . . . .	129
Database access . . . . .	130
Database security . . . . .	132
Databases vs. schemas . . . . .	133
Principals and schemas . . . . .	133
Default schemas . . . . .	134
Moving or copying a database . . . . .	134
Detaching and reattaching a database . . . . .	135
Backing up and restoring a database . . . . .	136

Understanding compatibility levels . . . . .	137
Conclusion . . . . .	138

**Chapter 4 Special databases 139**

System databases . . . . .	139
Understanding the master database . . . . .	139
Understanding the model database . . . . .	140
Introducing the tempdb database . . . . .	140
Understanding the resource database . . . . .	140
Understanding the msdb database . . . . .	141
Moving system databases . . . . .	142
Moving the master database . . . . .	143
The tempdb database . . . . .	144
Objects in tempdb . . . . .	144
Optimizations in tempdb . . . . .	146
Best practices . . . . .	147
tempdb contention . . . . .	148
tempdb space monitoring . . . . .	153
Database snapshots . . . . .	155
Creating a database snapshot . . . . .	156
Understanding space used by database snapshots . . . . .	159
Managing your snapshots . . . . .	161
Partially contained databases . . . . .	162
Configuring a contained database . . . . .	162
Creating contained users . . . . .	163
Understanding database collation changes . . . . .	166
Detecting uncontained features . . . . .	168
Conclusion . . . . .	169

**Chapter 5 Logging and recovery 171**

Transaction log internals . . . . .	171
Phases of recovery . . . . .	174
Page LSNs and recovery . . . . .	175
Log reading . . . . .	176



The log cache .....	177
Changes in log size .....	178
Understanding virtual log files .....	178
Maintaining a recoverable log .....	185
Automatically shrinking the log .....	187
Viewing the log file size .....	188
Database backup and restore .....	188
Understanding the types of backups .....	189
Understanding recovery models .....	190
Choosing a backup type .....	194
Restoring a database .....	195
Conclusion .....	201

## **Chapter 6 Table storage 203**

Table creation .....	203
Naming tables and columns .....	204
Avoiding reserved keywords .....	205
Using delimited identifiers .....	206
Understanding naming conventions .....	207
Choosing a data type .....	208
The NULL problem .....	233
User-defined data types .....	235
IDENTITY property .....	237
Sequence object .....	240
Internal storage .....	243
The sys.indexes catalog view .....	244
Data storage metadata .....	245
Catalog view queries .....	246
Data pages .....	248
The structure of data rows .....	257
How to find a physical page .....	259
Storage of fixed-length rows .....	262
Storage of variable-length rows .....	265
NULLS and variable-length columns .....	267

Storage of date and time data . . . . .	270
Storage of sql_variant data . . . . .	273
Constraints . . . . .	276
Constraint names and catalog view information . . . . .	277
Constraint failures in transactions and multiple-row data modifications . . . . .	278
Altering a table . . . . .	279
Changing a data type . . . . .	280
Adding a new column . . . . .	281
Adding, dropping, disabling, or enabling a constraint . . . . .	281
Dropping a column . . . . .	283
Internals of altering tables . . . . .	283
Heap modification internals . . . . .	286
Allocation structures . . . . .	286
Inserting rows . . . . .	288
Deleting rows . . . . .	288
Updating rows . . . . .	292
Conclusion . . . . .	295

**Chapter 7 Indexes: internals and management 297**

Overview . . . . .	298
SQL Server B-tree indexes . . . . .	299
Example 1: An index with a large key column . . . . .	300
Example 2: An index with a very narrow key column . . . . .	301
Tools for analyzing indexes . . . . .	302
Using the dm_db_index_physical_stats DMV . . . . .	302
Using sys.dm_db_database_page_allocations . . . . .	306
Understanding B-tree index structures . . . . .	308
Clustering key dependency . . . . .	308
Nonclustered B-tree indexes . . . . .	311
Constraints and indexes . . . . .	312
Index creation options . . . . .	313
IGNORE_DUP_KEY . . . . .	313
STATISTICS_NORECOMPUTE . . . . .	314

MAXDOP .....	314
Index placement .....	314
Physical index structures for B-trees .....	315
Index row formats .....	315
Clustered index structures .....	316
Non-leaf level(s) of a clustered index .....	317
Analyzing a clustered index structure .....	317
Nonclustered index structures .....	322
Indexes on computed columns and indexed views .....	333
SET options .....	333
Permissible functions .....	334
Schema binding .....	335
Indexes on computed columns .....	335
Implementation of a computed column .....	336
Persisted columns .....	337
Indexed views .....	338
Additional requirements .....	338
Creating an indexed view .....	339
Using an indexed view .....	340
Data modification internals .....	341
Inserting rows .....	342
Splitting pages .....	342
Deleting rows .....	346
Updating rows .....	354
Table-level vs. index-level data modification .....	357
Logging .....	358
Locking .....	358
Fragmentation .....	359
Managing B-tree index structures .....	360
Dropping indexes .....	360
Using the ALTER INDEX command .....	361
Detecting fragmentation .....	363
Removing fragmentation .....	364
Rebuilding an index .....	366
Online index building .....	367

Columnstore indexes .....	370
Creation of columnstore indexes.....	370
Storage of columnstore indexes .....	371
Columnstore index metadata .....	376
Conclusion .....	380

## **Chapter 8 Special storage 381**

Large object storage .....	381
Restricted-length large object data (row-overflow data).....	382
Unrestricted-length large object data .....	386
FILESTREAM and FileTable data .....	394
Enabling FILESTREAM data for SQL Server .....	395
Creating a FILESTREAM-enabled database.....	397
Creating a table to hold FILESTREAM data .....	397
Manipulating FILESTREAM data .....	399
Exploring metadata with FILESTREAM data .....	404
Creating a FileTable .....	406
Considering performance for FILESTREAM data .....	409
Summarizing FILESTREAM and FileTable .....	410
Sparse columns .....	411
Management of sparse columns .....	411
Column sets and sparse column manipulation.....	414
Physical storage .....	416
Metadata .....	419
Storage savings with sparse columns .....	420
Data compression .....	423
Vardecimal .....	423
Row compression.....	424
Page compression .....	433
Table and index partitioning .....	444
Partition functions and partition schemes.....	444
Metadata for partitioning.....	446
The sliding window benefits of partitioning .....	450

Partitioning a columnstore index . . . . .	452
Conclusion . . . . .	455

**Chapter 9 Special indexes 457**

Special indexes vs. ordinary indexes . . . . .	457
XML indexes. . . . .	458
Creating and maintaining XML indexes. . . . .	459
Using XQuery in SQL Server: internals . . . . .	463
Understanding how a query plan uses an XML index. . . . .	465
Using secondary XML indexes . . . . .	468
Working with XML indexes and schema-validated columns . . . . .	469
Using XML-specific information in query plans . . . . .	470
Spatial indexes. . . . .	471
Purpose of spatial indexes . . . . .	472
Composition of the spatial index. . . . .	475
How a spatial query uses a spatial index . . . . .	477
How to ensure that your spatial index is being used. . . . .	478
Spatial query plans and spatial indexes . . . . .	479
Nearest neighbor optimization in SQL Server 2012 . . . . .	481
Spatial index diagnostic stored procedures . . . . .	484
Diagnostics with the SQL Server 2012 spatial functions . . . . .	491
Full-text indexes . . . . .	492
Internal tables created by the full-text index . . . . .	494
Full-text index metadata views . . . . .	497
Full-text index creation . . . . .	498
Maintenance of a full-text index . . . . .	499
Full-text status metadata, configuration, and diagnostic information . . . . .	500
How a full-text index is used in a query. . . . .	501
A full-text query plan . . . . .	502
Extended event information for full-text queries. . . . .	503
Semantic indexes. . . . .	505
Conclusion . . . . .	511

## **Chapter 10 Query execution 513**

Introducing query processing and execution . . . . .	513
Iterators . . . . .	513
Properties of iterators . . . . .	515
Reading query plans . . . . .	517
Graphical plans . . . . .	517
Text plans . . . . .	518
XML plans . . . . .	518
Estimated vs. actual query plans . . . . .	518
Query plan display options . . . . .	520
Analyzing plans . . . . .	525
Scans and seeks . . . . .	526
Seekable predicates and covered columns . . . . .	528
Bookmark lookup . . . . .	531
Joins . . . . .	533
Aggregations . . . . .	545
Unions . . . . .	555
Advanced index operations . . . . .	560
Subqueries . . . . .	566
Parallelism . . . . .	580
Inserts, updates, and deletes . . . . .	598
Understanding data warehouses . . . . .	599
Using columnstore indexes and batch processing . . . . .	603
Adding new data . . . . .	607
Hints . . . . .	609
Conclusion . . . . .	610

## **Chapter 11 The Query Optimizer 611**

Overview . . . . .	611
Understanding the tree format . . . . .	612
Understanding optimization . . . . .	613
Search space and heuristics . . . . .	614
Rules . . . . .	614
Properties . . . . .	614

Storage of alternatives: the Memo .....	.617
Operators .....	.617
Optimizer architecture .....	.624
Before optimization .....	.625
Simplification .....	.625
Trivial plan/auto-parameterization .....	.625
Limitations .....	.627
The Memo: exploring multiple plans efficiently .....	.627
Statistics, cardinality estimation, and costing .....	.630
Statistics design .....	.631
Density/frequency information .....	.634
Filtered statistics .....	.636
String statistics .....	.637
Cardinality estimation details .....	.638
Limitations .....	.642
Costing .....	.643
Index selection .....	.645
Filtered indexes .....	.648
Indexed views .....	.649
Partitioned tables .....	.654
Partition-aligned index views .....	.658
Windowing functions .....	.658
Data warehousing .....	.659
Columnstore indexes .....	.660
Batch mode processing .....	.662
Plan shape .....	.667
Columnstore limitations and workarounds .....	.670
Updates .....	.670
Halloween Protection .....	.674
Split/Sort/Collapse .....	.674
Merge .....	.676
Wide update plans .....	.679
Non-updating updates .....	.681
Sparse column updates .....	.681

Partitioned updates . . . . .	682
Locking . . . . .	685
Partition-level lock escalation . . . . .	686
Distributed query . . . . .	687
Extended indexes . . . . .	689
Plan hinting . . . . .	689
Debugging plan issues . . . . .	691
{HASH   ORDER} GROUP . . . . .	692
{MERGE   HASH   CONCAT} UNION . . . . .	693
FORCE ORDER, {LOOP   MERGE   HASH} JOIN . . . . .	693
INDEX=<indexname>   <indexid> . . . . .	694
FORCESEEK . . . . .	695
FAST <number_rows> . . . . .	695
MAXDOP <N> . . . . .	696
OPTIMIZE FOR . . . . .	696
PARAMETERIZATION {SIMPLE   FORCED} . . . . .	698
NOEXPAND . . . . .	699
USE PLAN . . . . .	699
Hotfixes . . . . .	700
Conclusion . . . . .	701

## **Chapter 12 Plan caching and recompilation 703**

The plan cache . . . . .	703
Plan cache metadata . . . . .	704
Clearing plan cache . . . . .	704
Caching mechanisms . . . . .	705
Ad hoc query caching . . . . .	706
Optimizing for ad hoc workloads . . . . .	708
Simple parameterization . . . . .	711
Prepared queries . . . . .	717
Compiled objects . . . . .	719
Causes of recompilation . . . . .	722
Plan cache internals . . . . .	732
Cache stores . . . . .	732



Compiled plans . . . . .	734
Execution contexts . . . . .	734
Plan cache metadata . . . . .	735
Cache size management . . . . .	740
Costing of cache entries . . . . .	743
Objects in plan cache: the big picture . . . . .	744
Multiple plans in cache . . . . .	746
When to use stored procedures and other caching mechanisms . . . . .	747
Troubleshooting plan cache issues . . . . .	748
Optimization hints and plan guides . . . . .	752
Optimization hints . . . . .	752
Purpose of plan guides . . . . .	754
Types of plan guides . . . . .	755
Managing plan guides . . . . .	758
Plan guide considerations . . . . .	759
Conclusion . . . . .	764

**Chapter 13 Transactions and concurrency 765**

Concurrency models . . . . .	765
Pessimistic concurrency . . . . .	766
Optimistic concurrency . . . . .	766
Transaction processing . . . . .	766
ACID properties . . . . .	767
Transaction dependencies . . . . .	768
Isolation levels . . . . .	770
Locking . . . . .	774
Locking basics . . . . .	774
Spinlocks . . . . .	775
Lock types for user data . . . . .	775
Viewing locks . . . . .	786
Locking examples . . . . .	789
Lock compatibility . . . . .	794
Internal locking architecture . . . . .	796

Row-level locking vs. page-level locking . . . . .	803
Lock escalation . . . . .	804
Deadlocks . . . . .	806
Row versioning . . . . .	811
Row versioning details . . . . .	811
Snapshot-based isolation levels. . . . .	813
Choosing a concurrency model. . . . .	830
Controlling locking. . . . .	832
Lock hints. . . . .	832
Conclusion . . . . .	836

**Chapter 14 DBCC internals 837**

Shrinking files and databases . . . . .	837
Data file shrinking . . . . .	838
Log file shrinking . . . . .	840
DBCC SHRINKFILE . . . . .	840
AUTO_SHRINK . . . . .	841
Consistency checking . . . . .	841
Getting a consistent view of the database . . . . .	842
Processing the database efficiently. . . . .	845
Performing primitive system catalog consistency checks. . . . .	855
Performing allocation consistency checks. . . . .	856
Performing per-table logical consistency checks. . . . .	860
Processing columns . . . . .	866
Performing cross-table consistency checks. . . . .	881
Understanding DBCC CHECKDB output . . . . .	885
Reviewing DBCC CHECKDB options . . . . .	890
Performing database repairs . . . . .	893
Using consistency-checking commands other than DBCC CHECKDB . . . . .	898
Conclusion . . . . .	901

<i>Index</i> . . . . .	<i>903</i>
------------------------	------------