Module 3

Database structures

Module Overview

- Database structure internals
- Best practices for database files
- Best practices for tempdb

Lesson: Database Structure Internals

- Database file types
- Filegroups and files
- Pages and extents
- Shared vs uniform extents
- Page structure
- Page types
- Record types
- Allocation bitmaps and special pages
- Allocation units
- Investigating page structure

Database file types

- Data files
  - Divided into 8KB blocks called pages
  - Pages are grouped in 8, called extent
  - One primary file (mdf recommended)
  - X number of secondary files (ndf recommended)
  - More than one file can improve
    - Manageability
    - Performance (the disk subsystem might have limitations per file)
- Transaction log file
  - Written sequentially
  - No performance gain having more than one file
  - Initially buffered in memory, hardened at commit

Database Files and Filegroups
https://docs.microsoft.com/en-us/sql/relational-databases/databases/database-files-and-filegroups

SQL Server Transaction Log Architecture and Management Guide
https://docs.microsoft.com/en-us/sql/relational-databases/sql-server-transaction-log-architecture-and-management-guide

Database files types, continued

- Filestream
  - Option for varbinary(max) data type
  - More on this later
- Memory optimized
  - Use with memory-optimized tables
  - More on this in a later module

Filegroups and files

- Filegroups
  - Logical grouping of data files
  - Can bring administration and performance benefits
  - FILESTREAM and memory optimized tables has their own file groups
- Example:
  - Mandatory
  - Optional

Demo Create database

Pages and extents

- An extent is physically continuous pages
- Mixed/shared: shared between different database objects (*allocations* to be precise)
- Uniform: owned by a single database object (*allocation* to be precise)

0	1	2	3	4	5	6	7
8 Blue	9 Blue	10 Blue	11 Green	12 Green	13 Blue	14 Red	15 Green
16 Blue	17 Red	18 Blue	19 Blue	20 Green	21 Blue	22	23
24 Blue	25 Blue	26	27	28	29	30	31
32	33	…					

System

Shared

Shared

Uniform

Unallocated

Unallocated

Unallocated

Unallocated

Pages and Extents Architecture Guide
https://docs.microsoft.com/en-us/sql/relational-databases/pages-and-extents-architecture-guide

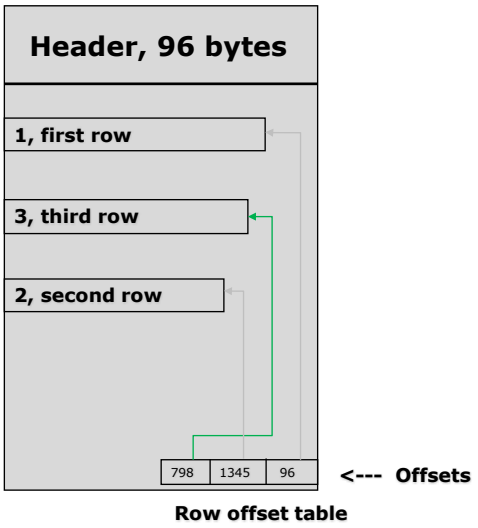
Shared vs uniform extents

- Traditionally, the first 8 pages allocated from shared extents
- As of SQL Server 2016, Uniform extents are used from 1:st page
  - Controlled by database setting:
  - SET MIXED_PAGE_ALLOCATION ON/OFF
  - This could be requested using trace flag 1118 prior to 2016
  - Trace flag 1118 has no effect as of 2016
  - Shared extents are still used for other things
    - Such as IAM pages
- SQL Server keep track of the first 8 pages and each extents that an allocation is using by the IAM page
  - Can be investigated using sys.dm_db_database_page_allocations()

Page Structure

Pages:

- 8 KB in size
- 96-byte page header



Inside the Storage Engine: Anatomy of a page
https://www.sqlskills.com/blogs/paul/inside-the-storage-engine-anatomy-of-a-page/

Understanding the Internals of a Data Page
https://www.sqlservercentral.com/articles/understanding-the-internals-of-a-data-page

Page Types

Pages:

- 8 KB in size
- 96-byte page header

	Page Types	
1	Data	Heap and clustered leaf pages
2	Index	Index pages (not including clustered leaf pages)
3, 4	LOB	Large value type pages, Row overflow pages
13	Boot	Page 9 in file 1
15	File Header	Page 0
11	PFS	More on this later
8	GAM	More on this later
9	SGAM	More on this later
10	IAM	More on this later
16	DIFF_MAP	More on this later
17	ML_MAP	More on this later

Pages and Extents Architecture Guide
https://docs.microsoft.com/en-us/sql/relational-databases/pages-and-extents-architecture-guide

SQL SERVER – Identifying Page Types
https://blog.sqlauthority.com/2016/04/16/sql-server-identifying-page-types/

SQL Server page types list
https://www.dabrowski.space/posts/sql-server-page-types-list/

Record types

- Record types:
  - Data records
    - Leaf pages for clustered indexes
    - Heap pages for heap tables
  - Index records
    - Leaf level index records
      - One can argue whether a *clustered* leaf page should be considered a data page or an index page. Best answer to that is probably "both".
    - Non-leaf level index records
  - Forwarding/forwarded records
  - Text records
  - Versioned records
  - Ghost records

SQL Server : Understanding the Data Record Structure
https://www.sqlservercentral.com/blogs/sql-serverunderstanding-the-data-record-structure

SQL SERVER – What is Forwarded Records and How to Fix Them?
https://blog.sqlauthority.com/2018/03/08/sql-server-forwarded-records-fix/

SQL Server Storage Engine: LOB Storage
https://aboutsqlserver.com/2013/11/05/sql-server-storage-engine-lob-storage/

Ghost cleanup process guide
https://docs.microsoft.com/en-us/sql/relational-databases/ghost-record-cleanup-process-guide

Allocation Bitmaps and Special Pages

- Page allocations are tracked by special pages:
- Global Allocation Map (GAM)
- Shared Global Allocation Map (SGAM)
- Differential Change Map (DCM)
- Bulk Change Map (BCM)
- Page Free Space (PFS)
  - Maps pages, one byte per page
  - Repeats approx. every 64 MB
- Index Allocation Map (IAM)
- All except File header and PFS:
  - One bit per extent
  - Repeats every 4GB

0 F hdr	1 PFS	2 GAM	3 SGAM	4	5	6 DCM	7 BCM
8 BlueIAM	9 Blue	10 Blue	11 GreenIAM	12 Green	13 Blue	14 RedIAM	15 Green
16 Blue	17 Red	18 Blue	19 Blue	20 Green	21 Blue	22	23
24 Blue	25 Blue	26	27	28	29	30	31
32	33	…					

Inside The Storage Engine: GAM, SGAM, PFS and other allocation maps
https://www.sqlskills.com/blogs/paul/inside-the-storage-engine-gam-sgam-pfs-and-other-allocation-maps/

SQL Server Extents, PFS, GAM, SGAM and IAM and related corruptions
https://techcommunity.microsoft.com/t5/sql-server-support-blog/sql-server-extents-pfs-gam-sgam-and-iam-and-related-corruptions/ba-p/1606011

Allocation Units

- Three types of allocation units
  - IN_ROW_DATA
  - LOB_DATA (optional)
  - ROW_OVERFLOW_DATA (optional)
- One of above per partition
- Also one per index

Demo LOB data

SQL Server – Understanding Allocation Units – In Row Data, LOB Data & Row Overflow Data
https://dataginger.com/2013/10/14/sql-server-understanding-allocation-units-in-row-data-lob-data-row-overflow-data/

sys.allocation_units (Transact-SQL)
https://docs.microsoft.com/en-us/sql/relational-databases/system-catalog-views/sys-allocation-units-transact-sql

Investigating page structure

- Decode the IAM page using:
  - sys.dm_db_database_page_allocations()
- Return the file/page/entry in offset table using
  - sys.fn_PhysLocFormatter(%%physloc%%)
    - Scalar function

    The three values combined in one column
  - sys.fn_PhysLocCracker(%%physloc%%)
    - Table function
    - Three columns
      - File #
      - Page #
      - Entry in offset table
    - Use CROSS APPLY to this
- Look at page internals using
  - DBCC PAGE
  - sys.dm_db_page_info

Demo Page structure

Use caution with sys.dm_db_database_page_allocations in SQL Server
https://www.mssqltips.com/sqlservertip/6309/use-caution-with-sysdmdbdatabasepageallocations-in-sql-server/

How to use the SQL Server sys.fn_PhysLocFormatter undocumented function
https://www.mssqltips.com/sqlservertip/1925/how-to-use-the-sql-server-sysfnphyslocformatter-undocumented-function/

SQL Servers Virtual Columns and Row Cracking
http://jongurgul.com/blog/sql-servers-virtual-columns-row-cracking/

Using DBCC PAGE to Examine SQL Server Table and Index Data
https://www.mssqltips.com/sqlservertip/1578/using-dbcc-page-to-examine-sql-server-table-and-index-data/

sys.dm_db_page_info (Transact-SQL)
https://docs.microsoft.com/en-us/sql/relational-databases/system-dynamic-management-views/sys-dm-db-page-info-transact-sql

Lesson: Best practices for database files

- Physical I/O vs. logical I/O
- Allocations within datafiles
- Instant File Initialization
- Autogrow and autoshrink
- Datafile shrinking
- Monitoring database files

Physical I/O vs. Logical I/O

- Logical I/O – pages read from memory
- Physical I/O – pages read from disk
- Buffer cache hit ratio – percentage of pages read from cache without having to be read from disk
  - Watch out: read-ahead reads are not included in this calculation
  - High value tell us nothing
  - Low value is a guarantee that we have a “lots of” physical I/O

Demo logical and physical IO

I/O vs Logical I/O
https://www.sqlservercentral.com/blogs/io-vs-logical-io

The Curious Case of… does SQL Server use a read/write thread per LUN?
https://www.sqlskills.com/blogs/paul/the-curious-case-of-does-sql-server-use-a-readwrite-thread-per-lun/

Reading Pages
https://docs.microsoft.com/en-us/sql/relational-databases/reading-pages

Sequential Read Ahead
https://techcommunity.microsoft.com/t5/sql-server-blog/sequential-read-ahead/ba-p/383480

Allocations within data files and autogrow

- Allocation is simple in a file group that has a single file
- In a multifile group, SQL Server uses:
  - Round robin allocation
  - Proportional fill
- Autogrow can result in different file sizes
  - One file larger than the other
  - Trace flag 1117 prior to 2016 cause all files to grow
  - *Autogrow all files* is default for tempdb
    - Cannot be changed
  - Filegroup setting for other databases
    - Default is OFF

SQL Server Proportional Fill Algorithm Example
https://www.mssqltips.com/sqlservertip/4890/sql-server-proportional-fill-algorithm-example/

Investigating the proportional fill algorithm
https://www.sqlskills.com/blogs/paul/investigating-the-proportional-fill-algorithm/

Expand All Database Files Simultaneously Using SQL Server 2016 AUTOGROW_ALL_FILES
https://www.mssqltips.com/sqlservertip/4937/expand-all-database-files-simultaneously-using-sql-server-2016-autogrowallfiles/

ALTER DATABASE (Transact-SQL) File and Filegroup Options
https://docs.microsoft.com/en-us/sql/t-sql/statements/alter-database-transact-sql-file-and-filegroup-options

Instant File Initialization

- Improves performance by skipping zeroing of data pages on file creation and growth
- Does not apply for log files
- Disabled by default
  - Potential security consideration: you can read data from deleted files using DBCC PAGE
- Enable if above is OK
- Check
  - See boot information in errorlog file
  - sys.dm_server_services.instant_file_initialization_enabled olumn
- Checkbox for this in SQL Server setup
- Change using “Perform volume maintenance tasks” security policy
  - Assign to SQL Server service account

Demo Instant File Initialization

Database Instant File Initialization
https://docs.microsoft.com/en-us/sql/relational-databases/databases/database-instant-file-initialization

Instant File Initialization
https://www.brentozar.com/blitz/instant-file-initialization/

Autogrow

- File property
- Autogrow
  - Automatically expands a file when space is low
  - Manually manage file size for better performance
  - Auto grow best practices:
    - Leave enabled to avoid downtime in case file size limit is reached
    - Set to a fixed rather than a percentage

Considerations for the autogrow and autoshrink settings in SQL Server
https://docs.microsoft.com/en-us/troubleshoot/sql/admin/considerations-autogrow-autoshrink

Adjust autogrow values for database files
https://karaszi.com/adjust-autogrow-values-for-database-files

Datafile shrinking

- Shrink
  - Causes index fragmentation
  - Resource intensive
  - Degrades performance
- Best used for
  - Emptying a file before removing it
  - Changing a file or file group to read only
  - Reclaiming free space after deleting a ***large amount*** of data
- Alternative
  - Move all indexes to a new file group
  - Move heaps by creating clustered index and then possibly dropping it
  - Remove old file group
- Things that can cause shrink to take a very long time
  - LOB data (a table scan for each LOB page moved)
  - Heaps (update all non-clustered indexes for every row on every data-page that was moved)
  - Blocking (shrink waits forever)

Demo Do not shrink

Why you want to be restrictive with shrink of database files
https://karaszi.com/why-you-want-to-be-restrictive-with-shrink-of-database-files

Why you should not shrink your data files
https://www.sqlskills.com/blogs/paul/why-you-should-not-shrink-your-data-files/

Autoshrink

- Database setting
- Just say No
  - Verify that no database has this
  - sys.databases
    - is_auto_shrink_on

Considerations for the autogrow and autoshrink settings in SQL Server
https://docs.microsoft.com/en-us/troubleshoot/sql/admin/considerations-autogrow-autoshrink

Auto-Shrink Enabled or Job has Shrink Steps
https://www.brentozar.com/blitz/auto-shrink-enabled/

SHRINK a data file? Just say NO!!!
https://www.theboreddba.com/Categories/indexes/SHRINK-a-data-file-Just-say-NO.aspx

Auto-shrink – turn it OFF!
https://www.sqlskills.com/blogs/paul/auto-shrink-turn-it-off/

Monitoring database files

- Database file configuration
  - SSMS database properties
  - sys.database_files and sys.filegroups
  - sys.master_files
  - sys.dm_db_file_space_usage
- Database file activity
  - SSMS Activity Monitor – Data File I/O
  - Wait statistics – PAGEIOLATCH_*, WRITELOG
  - sys.dm_io_virtual_file_stats

sys.database_files (Transact-SQL)
https://docs.microsoft.com/en-us/sql/relational-databases/system-catalog-views/sys-database-files-transact-sql

sys.filegroups (Transact-SQL)
https://docs.microsoft.com/en-us/sql/relational-databases/system-catalog-views/sys-filegroups-transact-sql

sys.master_files (Transact-SQL)
https://docs.microsoft.com/en-us/sql/relational-databases/system-catalog-views/sys-master-files-transact-sql

sys.dm_io_virtual_file_stats (Transact-SQL)
https://docs.microsoft.com/en-us/sql/relational-databases/system-dynamic-management-views/sys-dm-io-virtual-file-stats-transact-sql

What Virtual Filestats Do, and Do Not, Tell You About I/O Latency
https://sqlperformance.com/2013/10/t-sql-queries/io-latency

Lesson: Best practices for tempdb

- Tempdb usage
- Tempdb configuration

Tempdb Usage

- Stores:
  - Internal objects
  - Version store, which includes
    - Trigger's *deleted* and *inserted* tables
    - Working space for online index operations
  - Some user objects, for instance
    - Temp tables
    - Table variables
- Recreated/reinitialized at startup
- You don't want tempdb to run out of space

tempdb database
https://docs.microsoft.com/en-us/sql/relational-databases/databases/tempdb-database

Tempdb Configuration

- Tempdb configuration
  - Enable instant file initialization
  - Size for expected workload
  - Consider to isolate tempdb data files for management purposes
  - Use multiple data files
    - 1 per CPU core up to 8 cores as a starting point
  - Enable auto growth as a contingency
  - Make files grow equally
    - Taken care of for you since 2016
  - Consider memory optimized tempdb metadata in 2019
    - Server-level setting
    - Makes 12 system tables use non-durable memory-optimized table
- Further improvements in 2022
  - Concurrent updates to GAM and SGAM pages

Demo Tempdb

Cheat Sheet: How to Configure TempDB for Microsoft SQL Server
https://www.brentozar.com/archive/2016/01/cheat-sheet-how-to-configure-tempdb-for-microsoft-sql-server/

Improve scalability with system page latch concurrency enhancements in SQL Server 2022
https://cloudblogs.microsoft.com/sqlserver/2022/07/21/improve-scalability-with-system-page-latch-concurrency-enhancements-in-sql-server-2022/

Lab 3: Database structures

- Ex 1: Verifying Instant File Initialization
- Ex 2: Verifying number of tempdb datafiles

Estimated Time: 30 minutes