Microsoft

Files & Filegroups

Agenda

- File & Filegroups
- Design Considerations

Files & Filegroups

SQL Server File Structure

Data Files contain data and objects such as tables, indexes, stored procedures, and views.

Data files can be of two types: Primary or Secondary.

- Primary Data File, which contains:
  - Startup information for the database and points to other files in the database.
  - User data and objects can be stored in this file
- Secondary data files are optional and can be used to spread data across multiple files/disks by putting each file on a different disk drive.

Database Files

File	Description
Primary (mdf)	Contains startup information for the database and points to the other files in the database. Every database has one primary data file.
Secondary (ndf)	Optional user-defined data files. Data can be spread across multiple disks by putting each file on a different disk drive.
Transaction Log (ldf)	The log holds information used to recover the database. There must be at least one log file for each database.

Logical and physical file names

SQL Server files have two **file name types**:

logical_file_name: The logical_file_name is the name used to refer to the physical file in all Transact-SQL statements.

os_file_name: The os_file_name is the name of the physical file including the directory path.

Filegroups

The **primary filegroup** contains the primary data file and any secondary files that aren't put into other filegroups.

User-defined filegroups can be created to group data files together for administrative, data allocation, and placement purposes.

Queries for data from the table will be spread across the three disks

- It will improve performance.
- Files and filegroups let you easily add new files to new disks.

Files & Filegroups Fill Strategy

Data from a table is spread across all the files in the filegroups (remember extents and pages)

SQL Server uses a proportional fill algorithm to distribute data across those files.

If **file1.ndf** has 100MB, and **file2.ndf** has 200MB

Then:

- For every extent in file1.ndf
- 2 extents are allocated in file2.ndf
- Both files will get filled at almost the same time!

Rules for designing files and filegroups

- A file or filegroup cannot be used by more than one database.
  - For example, file **sales.mdf** and **sales.ndf**, which contain data and objects from the sales database, can't be used by any other database.

A file can be a member of only one filegroup.

Transaction log files are never part of any filegroups.

Database Creation with Files & Filegroups

```
USE master;
GO
CREATE DATABASE MyDB
ON PRIMARY
  ( NAME='MyDB_Primary',
    FILENAME=
       ‘var/opt/mssql/data/MyDB_PRI.mdf',
    SIZE=4MB,
    MAXSIZE=10MB,
    FILEGROWTH=1MB),
```

Database Creation with Files & Filegroups

```
FILEGROUP MyDB_FG1
  ( NAME = 'MyDB_FG1_Dat1',
    FILENAME = ‘var/opt/mssql/data/MyDB_FG1_1.ndf',
    SIZE = 1MB,
    MAXSIZE=10MB,
    FILEGROWTH=1MB),
  ( NAME = 'MyDB_FG1_Dat2',
    FILENAME = ‘var/opt/mssql/data/MyDB_FG1_2.ndf',
    SIZE = 1MB,
    MAXSIZE=10MB,
    FILEGROWTH=1MB),
```

Database Creation with Files & Filegroups

```
LOG ON
  ( NAME='MyDB_log',
    FILENAME = ‘/var/opt/mssql/data/MyDB.ldf',
    SIZE=1MB,
    MAXSIZE=10MB,
    FILEGROWTH=1MB);
GO
```

Database Creation with Files & Filegroups

```
ALTER DATABASE MyDB
  MODIFY FILEGROUP MyDB_FG1 DEFAULT;
GO

USE MyDB;
CREATE TABLE MyTable
  ( col_a int PRIMARY KEY,
    col_b char(8) )
ON MyDB_FG1;
GO
```

Files & Filegroups Design Guidelines

Files & Filegroups Design Guidelines

Most databases will work well with a single data file and a single transaction log file.

If you use multiple data files, create a second filegroup and make that filegroup the default filegroup

To maximize performance, create files or filegroups on different available disks as possible

Files & Filegroups Design Guidelines

Use filegroups to enable placement of objects on specific physical disks.

Put tables used in the same join queries in different filegroups. This step will improve performance, because of parallel disk I/O searching for joined data.

Don't put the transaction log file(s) on the same physical disk that has the other files and filegroups.