

SQL SERVER

知识点记录

每天学一点，每天进步一点

目录

[SQL Sever中数据库文件的存放方式，文件和文件组 3](#_Toc3061)

[简介 3](#_Toc1012)

[理解文件和文件组 3](#_Toc28774)

[为什么通过文件组来管理文件 3](#_Toc1328)

[文件的分类 4](#_Toc29211)

[创建和使用文件组 5](#_Toc15476)

[使用多个文件的优点和缺点 5](#_Toc12625)

[总结 6](#_Toc369)

[SQL SERVER中的页和区体系结构 7](#_Toc17817)

[页和区 7](#_Toc22909)

[页 7](#_Toc17133)

[下表说明了SQL SERVER数据库的数据文件中所使用的页类型。 7](#_Toc1206)

[大型行支持 8](#_Toc15550)

[区 9](#_Toc30066)

# SQL Sever中数据库文件的存放方式，文件和文件组

## 简介

在SQL SERVER中，数据库在硬盘上的存储方式和普通文件在Windows中的存储方式没有什么不同，仅仅是几个文件而已。

SQL SERVER通过管理逻辑上的文件组的方式来管理文件。

理解文件和文件组的概念对于更好的配置数据库来说是最基本的知识。

## 理解文件和文件组

在SQL SERVER中，通过文件组这个逻辑对象存放数据的文件进行管理。

我们看到的逻辑数据库由一个或者多个文件组构成，

而文件组管理着磁盘上的文件，文件中存放着SQL SERVER的实际数据

## 为什么通过文件组来管理文件

对于用户角度来说，需对创建的对象指定存储的文件只有三种数据对象：表、索引和大对象(LOB)

使用文件组可以隔离用户和文件，使得用户针对文件组来建立表和索引，而不是实际磁盘中的文件。当文件移动或修改时，由于用户建立的表和索引是建立在文件组上的，并不依赖具体文件，这大大加强了可管理性。

还有一点是，使用文件组来管理文件可以使得同一文件组内的不同文件分布在不同的硬盘中，极大的提高了IO性能。

SQL SERVER会根据每个文件设置的初始大小和增长量自动分配新加入的空间，假设在同一文件组中的文件A设置的大小为文件B的两倍，新增一个数据占用三页(Page)，则按比例将2页分配到文件A中，1页分配到文件B中。

## 文件的分类

首要文件(.mdf)：这个文件是必须有的，而且只能有一个。这个文件额外存放了其他文件的位置等信息。

次要文件(.ndf)：可以建任意多个，用于不同目的的存放。

日志文件(.ldf)，存放日志。

上面说的各种文件名扩展名可以任意修改，推荐使用默认的扩展名。在SQL SERVER 2008之后，还新增了文件流数据文件和全文索引文件。

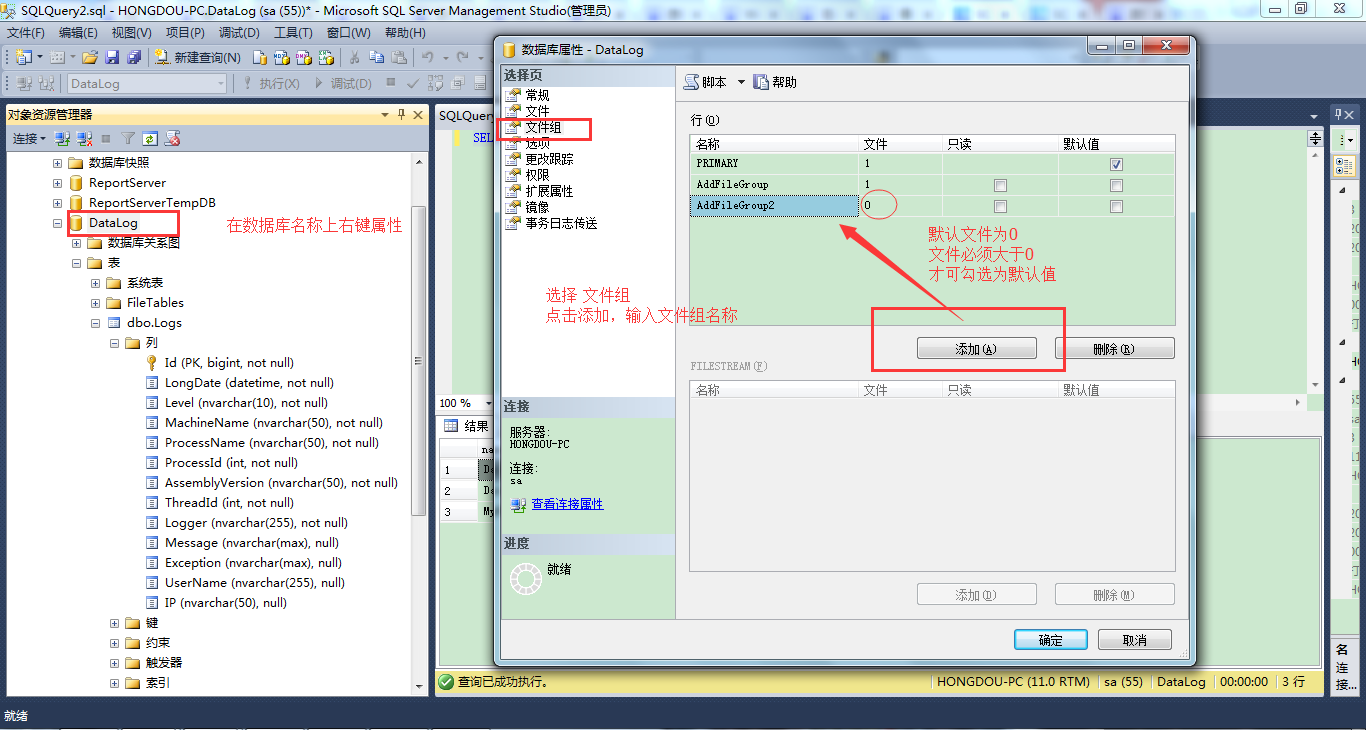
通过sys.database\_files系统试图可以查看SQL SERVER数据库中文件的情况。

SELECT name,type\_desc,physical\_name,state\_desc,size,growth FROM sys.database\_files

注意type\_desc列，首要文件和次要文件都是ROWS，表示存储的是行(ROW)

如果一个表是存在物理上的多个文件中时，则表的数据页的组织为N(N为具体的几个文件)个B树。而不是一个对象为一个B树。

## 创建和使用文件组



文件也可以用上面的方法创建，并可以指定文件组为新增文件组

注意：文件的所属文件组一旦设置了就不能够更改了。

下面给出一个创建索引时，指定存放到文件组的SQL实例：

--添加索引时，指定存放的文件组

　　CREATE NONCLUSTERED INDEX IX\_Logs\_LongDate on Logs(LongDate) ON AddFileGroup --红色部分是主角，指定文件组。

　　--如果不指定文件组，则存放到默认的文件组当中

　　CREATE NONCLUSTERED INDEX IX\_Logs\_LongDate on Logs(LongDate)

## 使用多个文件的优点和缺点

在大多数情况下，小型的数据库并不需要创建多个文件来存放数据。但是随着数据的增长，单个文件的弊端就会出现。

首先，使用多个文件分布到不同的磁盘分区(多个硬盘)能够极大提高IO性能。

其次，多个文件对于数据比较多的数据库来说，备份和恢复都会方便。

但是，多文件需要占用更多的磁盘空间，因为每个文件中都有自己的一套B树组织方式和自己的增长空间。当然也有自己的碎片。

总体来说，多个文件带来的优点是远远大于弊端的。

## 总结

 这里对SQL SERVER中文件和文件组的概念进行了简单阐述，并在文中讲述了文件和文件组的配置方式。按照业务组织好不同的文件组来分布不同的文件，以提升性能。

# SQL SERVER中的页和区体系结构

## 页和区

SQL SERVER中数据存储的基本单位是页。为数据库中的数据文件(.mdf或.ndf)分配的磁盘空间可以从逻辑上划分成页(从0到n连续编号)。磁盘I/O操作在页级执行。也就是说，SQL SERVER读取或写入所有数据页。

区是八个物理上连续的页的集合，用来有效地管理页。所有页都存储在区中。

## 页

在SQL SERVER中，页的大小为8KB。这意味着SQL SERVER数据库中每MB有128页。每页的开头是96字节的表头，用于存储有关页的系统信息。此信息包括页码、页类型、页的可用空间以及拥有该页的对象的分配单元ID。

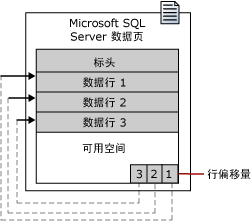
### 下表说明了SQL SERVER数据库的数据文件中所使用的页类型。



注意：日志文件不包含页，而是包含一系列日志记录

在数据页上，数据行紧接着标头按顺序放置。页的末尾是行偏移表，对于页中的每一行，每个行偏移表都包含一个条目。每个条目记录对应行的第一个字节与页首的距离。行偏移表中的条目的顺序与页中行的顺序相反。

行偏移量的大小是2字节。



### 大型行支持

行不能跨页，但是行的部分可以移出行所在的页，因此行实际可能非常大。页的单个行中的最大数据量和开销是8,060字节(8KB)。但是，这不包括用Tex/Image页类型存储的数据。包含**varchar**、**nvarchar**、**varbinary**或**sql\_variant**列的表不受此限制的约束。当表中的所有固定列和可变列的行的总大小超过限制的8,060字节时，SQL SERVER将从最大长度的列开始动态将一个或多个可变长度列移动到ROW\_OVERFLOW\_DATA分配单元中的页。每当插入或更新操作将行的总大小增大到超过限制的8,060字节时，将会执行此操作。将列移动到ROW\_OVERFLOW\_DATA分配单元中的页后，将在IN\_ROW\_DATA分配单元中的原始页中的原始页维护24字节的指针。如果后续操作减小了行的大小，SQL SERVER会动态将列移回到原始数据页。

## 区

区是管理空间的基本单位。一个区是八个物理上连续的页(即64KB)。这意味着SQL SERVER数据库中每MB有16个区。

为了使空间分配更有效，SQL SERVER不会将所有区分配给包含少量数据的表。SQL SERVER有两种类型的区：

统一区，由单个对象所有。区中的所有8页只能由所属对象使用。

混合区，最多可由八个对象共享。区中八页的每页可由不同的对象所有。

通常从混合区向新表或索引分配页。当表或索引增长到8页时，将变成使用统一区进行后续分配。如果对现有表创建索引，并且该表包含的行足以在索引中生成8页，则对该索引的所有分配都使用统一区进行。

