

SQLServer2005对海量数据处理 PDF转换可能丢失图片或格式
，建议阅读原文

https://www.100test.com/kao_ti2020/251/2021_2022_SQLServer2_c97_251409.htm 超大型数据库的大小常常达到数百GB，有时甚至要用TB来计算。而单表的数据量往往会达到上亿的记录，并且记录数会随着时间而增长。这不但影响着数据库的运行效率，也增大数据库的维护难度。除了表的数据量外，对表不同的访问模式也可能会影响性能和可用性。这些问题都可以通过对大表进行合理分区得到很大的改善。当表和索引变得非常大时，分区可以将数据分为更小、更容易管理的部分来提高系统的运行效率。如果系统有多个CPU或是多个磁盘子系统，可以通过并行操作获得更好的性能。所以对大表进行分区是处理海量数据的一种十分高效的方法。本文通过一个具体实例，介绍如何创建和修改分区表，以及如何查看分区表。

1 SQL Server 2005 SQL Server 2005是微软在推出SQL Server 2000后时隔五年推出的一个数据库平台，它的数据库引擎为关系型数据和结构化数据提供了更安全可靠存储功能，使用户可以构建和管理用于业务的高可用和高性能的数据应用程序。此外SQL Server 2005结合了分析、报表、集成和通知功能。这使企业可以构建和部署经济有效的BI解决方案，帮助团队通过记分卡、Dashboard、Web Services和移动设备将数据应用推向业务的各个领域。无论是开发人员、数据库管理员、信息工作者还是决策者，SQL Server 2005都可以提供出创新的解决方案，并可从数据中获得更多的益处。它所带来的新特性，如T-SQL的增强、数据分区、服务代理和与.Net Framework的集成等，在易管理性、可用性、可伸缩性和安全

性等方面都有很大的增强。2 表分区的具体实现方法 表分区分为水平分区和垂直分区。水平分区将表分为多个表。每个表包含的列数相同，但是行更少。例如，可以将一个包含十亿行的表水平分区成 12 个表，每个小表表示特定年份内一个月的数据。任何需要特定月份数据的查询只需引用相应月份的表。而垂直分区则是将原始表分成多个只包含较少列的表。水平分区是最常用分区方式，本文以水平分区来介绍具体实现方法。水平分区常用的方法是根据时期和使用对数据进行水平分区。例如本文例子，一个短信发送记录表包含最近一年的数据，但是只定期访问本季度的数据。在这种情况下，可考虑将数据分成四个区，每个区只包含一个季度的数据。

2.1 创建文件组

建立分区表先要创建文件组，而创建多个文件组主要是为了获得好的 I/O 平衡。一般情况下，文件组数最好与分区数相同，并且这些文件组通常位于不同的磁盘上。每个文件组可以由一个或多个文件构成，而每个分区必须映射到一个文件组。一个文件组可以由多个分区使用。为了更好地管理数据(例如，为了获得更精确的备份控制)，对分区表应进行设计，以便只有相关数据或逻辑分组的数据位于同一个文件组中。使用 ALTER DATABASE，添加逻辑文件组名：ALTER DATABASE [DeanDB] ADD FILEGROUP [FG1] DeanDB为数据库名称，FG1文件组名。创建文件组后，再使用 ALTER DATABASE 将文件添加到该文件组中：ALTER DATABASE [DeanDB] ADD FILE (NAME = NFG1, FILENAME = NC:DeanDataFG1.ndf, SIZE = 3072KB, FILEGROWTH = 1024KB) TO FILEGROUP [FG1] 类似的建立四个文件和文件组，并把每一个存储数据的文件放在不同的磁盘驱动器里。

2.2 创建分区函数 创建分区表必须先确定分区的功能机制，表进行分区的标准是通过分区函数来决定的。创建数据分区函数有RANGE “ LEFT | / RIGHT ” 两种选择。代表每个边界值在局部的哪一边。例如存在四个分区，则定义三个边界点值，并指定每个值是第一个分区的上边界 (LEFT) 还是第二个分区下边界 (RIGHT)[1]。代码如下：CREATE PARTITION FUNCTION [SendSMSPF](datetime) AS RANGE RIGHT FOR VALUES (20070401, 20070701, 20071001)

2.3 创建分区方案 创建分区函数后，必须将其与分区方案相关联，以便将分区指向至特定的文件组。就是定义实际存放数据的媒体与各数据块的对应关系。多个数据表可以共用相同的数据分区函数，一般不共用相同的数据分区方案。可以通过不同的分区方案，使用相同的分区函数，使不同的数据表有相同的分区条件，但存放在不同的媒介上。创建分区方案的代码如下：

CREATE PARTITION SCHEME [SendSMSPS] AS PARTITION [SendSMSPF] TO ([FG1], [FG2], [FG3], [FG4])

2.4 创建分区表 建立好分区函数和分区方案后，就可以创建分区表了。分区表是通过定义分区键值和分区方案相联系的。插入记录时，SQL Server会根据分区键值的不同，通过分区函数的定义将数据放到相应的分区。从而把分区函数、分区方案和分区表三者有机的结合起来。创建分区表的代码如下：CREATE TABLE SendSMSLog ([ID] [int] IDENTITY(1,1) NOT NULL, [IDNum] [nvarchar](50) NULL, [SendContent] [text] NULL [SendDate] [datetime] NOT NULL,) ON SendSMSPS(SendDate)

100Test 下载频道开通，各类考试题目直接下载。详细请访问 www.100test.com