Atitit **聚集索引 页分裂**

**9、** **聚集索引没有建在表的顺序字段上，该表容易发生页分裂**

　　比如订单表，有订单编号orderid，也有客户编号contactid，那么聚集索引应该加在哪个字段上呢？对于该表，订单编号是顺序添加的，如果在orderid上加聚集索引，新增的行都是添加在末尾，这样不容易经常产生页分裂。然而，由于大多数查询都是根据客户编号来查的，因此，将聚集索引加在contactid上才有意义。而contactid对于订单表而言，并非顺序字段。

　　比如“张三”的“contactid”是001，那么“张三”的订单信息必须都放在这张表的第一个数据页上，如果今天“张三”新下了一个订单，那该订单信息不能放在表的最后一页，而是第一页！如果第一页放满了呢？很抱歉，该表所有数据都要往后移动为这条记录腾地方。

　　SQL Server的索引和Oracle的索引是不同的，SQL Server的聚集索引实际上是对表按照聚集索引字段的顺序进行了排序，相当于oracle的索引组织表。SQL Server的聚集索引就是表本身的一种组织形式，所以它的效率是非常高的。也正因为此，插入一条记录，它的位置不是随便放的，而是要按照顺序放在该放的数据页，如果那个数据页没有空间了，就引起了页分裂。所以很显然，聚集索引没有建在表的顺序字段上，该表容易发生页分裂。

　　曾经碰到过一个情况，一位哥们的某张表重建索引后，插入的效率大幅下降了。估计情况大概是这样的。该表的聚集索引可能没有建在表的顺序字段上，该表经常被归档，所以该表的数据是以一种稀疏状态存在的。比如张三下过20张订单，而最近3个月的订单只有5张，归档策略是保留3个月数据，那么张三过去的 15张订单已经被归档，留下15个空位，可以在insert发生时重新被利用。在这种情况下由于有空位可以利用，就不会发生页分裂。但是查询性能会比较低，因为查询时必须扫描那些没有数据的空位。

　　重建聚集索引后情况改变了，因为重建聚集索引就是把表中的数据重新排列一遍，原来的空位没有了，而页的填充率又很高，插入数据经常要发生页分裂，所以性能大幅下降。

　　对于聚集索引没有建在顺序字段上的表，是否要给与比较低的页填充率？是否要避免重建聚集索引？是一个值得考虑的问题！