<http://blog.csdn.net/miracleww/article/details/53352738>

1. 什么是索引：

索引就像是书的目录，是与表或视图关联的磁盘上结构，可以加快从表或视图中检索行的速度。索引中包含由表或视图中的一列或多列生成的键。这些键存储在一个结构（BTree）中，使SQL可以快速有效地查找与键值关联的行。

2. 为什么要建立索引，即索引的优点：

①  建立索引的列可以保证行的唯一性，生成唯一的rowId

②  建立索引可以有效缩短数据的检索时间

③  建立索引可以加快表与表之间的连接

④  为用来排序或者是分组的字段添加索引可以加快分组和排序顺序

3. 索引的缺点：

①  创建索引和维护索引需要时间成本，这个成本随着数据量的增加而加大

②  创建索引和维护索引需要空间成本，每一条索引都要占据数据库的物理存储空间，数据量越大，占用空间也越大（数据表占据的是数据库的数据空间）

③  会降低表的增删改的效率，因为每次增删改索引需要进行动态维护，导致时间变长

4. 什么样的表跟列要建立索引：

①  总的来说就是数据量大的，经常进行查询操作的表要建立索引

②  表中字段建立索引应该遵循几个原则：

1)        越小的数据类型通常更好：越小的数据类型通常在磁盘、内存中都需要更少的空间，处理起来更快。

2)          简单的数据类型更好：整型数据比起字符，处理开销更小，因为字符串的比较更复杂，处理起来也更耗时。

3)        尽量避免NULL：应该指定列为NOT NULL。含有空值的列很难进行查询优化，因为它们使得索引、索引的统计信息以及比较运算更加复杂。你应该用0、一个特殊的值或者一个空串代替空值。

4)        对非唯一的字段，例如“性别”这种大量重复值的字段，增加索引也没有什么意义，所以索引的建立应当更多的选取唯一性更高的字段。

③  表与表连接用于多表联合查询的约束条件的字段应当建立索引

④  用于排序的字段可以添加索引，用于分组的字段应当视情况看是否需要添加索引。

⑤  添加多列索引的时候，对应的多条件查询可以触发该索引的同时，索引最左侧的列的单条件查询也可以触发。

⑥  如果有些表注定只会进行查询所有，也就没必要添加索引，因为查询全部只能进行全量搜索即扫描全表。

5. 索引的原理：

索引的原理大致概括为以空间换时间，数据库在未添加索引的时候进行查询默认的是进行全量搜索，也就是进行全局扫描，有多少条数据就要进行多少次查询，然后找到相匹配的数据就把他放到结果集中，直到全表扫描完。而建立索引之后，会将建立索引的KEY值放在一个n叉树上（BTree）。因为B树的特点就是适合在磁盘等直接存储设备上组织动态查找表，每次以索引进行条件查询时，会去树上根据key值直接进行搜索，次数约为log总条数，底数为页面存储数，例如一个100万数据的表，页面存储数为100，那么有索引的查询次数为3次log1000000100，但是全量搜索为100万次搜索，这种方式类似于二分法，但是这个是n分法。

索引对增删改的影响实际数据修改测试：

一个表有字段A、B、C，同时进行插入10000行记录测试

在没有建索引时平均完成时间是2.9秒

在对A字段建索引后平均完成时间是6.7秒

在对A字段和B字段建索引后平均完成时间是10.3秒

在对A字段、B字段和C字段都建索引后平均完成时间是11.7秒

从以上测试结果可以明显看出索引对数据修改产生的影响

版权声明：本文为博主原创文章，未经博主允许不得转载。