

96 MySQL是如何根据成本优化选择执行计划的？（下）

今天是我们讲解根据成本优化选择执行计划的最后一讲，下周就要给大家讲解基于规则的执行计划优化了，也就是MySQL是如何自动调整我们的SQL语句为性能比较优化的方式，好，那今天一起看看多表关联查询是如何选择执行计划的。

其实多表查询的执行计划选择思路，基本跟单表查询的执行计划选择思路是类似的，因为大家应该都记得，单表查询的时候，主要就是对这个表的多种访问方式（全表查询，各个索引查询）来根据一定的公式计算出来每种访问方式的成本，接着选择一个成本最低的访问方式，那么就可以确定下来这个表怎么访问了。

可能有的人看了之前的两讲，会觉得似乎这种成本计算的方式也不是太靠谱，因为里面有些过程感觉怪
怪的，不过这个没办法，其实即使让你来设计，也很难设计出完全公平、完全精准的成本预估算法来

因为要在一个查询执行之前，就可以针对不同的访问方法精准计算他的成本，那是根本不现实的，最后只能是根据一些相对较为简单粗暴的办法，大致估算一下，估算结果可能不是太准确，但是也没办法了，反正算出来也就这么比较就是了。

那么接着如果我们要看看多表关联的成本计算访问和执行计划选择方式，那就很简单了，因为大家应该还记得，多表关联的语句，比如：

```
select * from t1 join t2 on t1.x1=t2.x1 where t1.x2=xxx and t1.x3=xxx and t2.x4=xxx and t2.x5=xxx
```

就这么一个语句，大家应该还记得他里面的访问过程

一般来说，都会先选择一个驱动表，比如t1作为驱动表，此时就需要根据t1.x2=xxx和t1.x3=xxx这个条件从表里查询一波符合条件的数据出来，此时就有一个问题了，这里用到了t1的两个字段来筛选数据，可能x2和x3字段都建了索引了，此时到底选择哪个索引呢？或者干脆直接就是全表扫描？

此时就会按照之前讲的那套方法来计算针对t1表查询的全表扫描和不同索引的成本，选择一个针对t1表的最佳访问方式，用最低成本从t1表里查出符合条件的数据来，接着就根据这波数据得去t2表里查数据，按照连接条件t1.x1=t2.x1去查，同时要符合t2.x4=xxx和t2.x5=xxx这两个条件。

此时一样会根据之前讲解的办法去估算，针对t2表的全表扫描以及基于x4、x5、x1几个字段不同索引的访问的成本，挑选一个成本最低的方法，然后从t2表里把数据给查找出来，就可以，这就完成了多表关联！

所以大家可以看到，其实多表关联的成本估算以及执行计划选择方式，跟单表关联基本上是差不多的，只不过多表关联要多查几个表罢了。

End