

➤ 查询优化的准则

分布式查询优化的准则是使通信费用最低和响应时间最短，即以最小的总代价，在最短的响应时间内获得需要的数据。所谓响应时间，是从接收查询到完成查询所需的时间，它既与通信时间有关，也与局部处理时间有关，而通信费用与所传输的数据量和通信数成正比。

➤ 查询代价分析

通信费用也与分布式数据库系统所基于的通信网络的类型有关，对不同的通信网络类型有不同的查询处理优化算法。

(1) 在远程通信网络中

在远程通信网络中，如果受到网络传输带宽的限制（如 10Kbit/s），也就是说信息传送率的数量级低于磁盘与主存储器之间的传输率时，分布式查询时查询的局部处理时间与通信所需时间相比，可以忽略不计，减少通信费用成为分布式查询优化的主要目标，而通信费用与所传输的数据量成正比。因此，在基于远程通信网的分布式数据库系统的查询处理中，常常以减少传输的次数和数据量作为优化的重要目标。

(2) 在高速局域网中

在这种情况下，传输时间比局部处理时间要短得多，往往以响应时间作为优化目标。响应时间既与通信时间有关，也与局部处理时间有关，但局部处理时间是关键，所以减少局部处理的时间是问题的主要方面。

在某些情况下，查询处理同时以减少通信费用与响应时间作为优化目标。这时，算法往往需要在这两者之间做出权衡。另外，与集中式查询处理不同，为了减少局部处理时间，除了要考虑 CPU 和 I/O 时间外，在分布式查询中还应考虑如何进行并行查询处理的问题，还需要随时了解各个站点上的数据变化情况，以便估算各种查询策略的代价。

(3) 查询代价的估算方法

对于分布式数据库环境，除估算输入/输出 (I/O) 操作次数，以及 CPU 的使用情况外，还必须考虑数据在不同场地间的传输代价。场地间的传输代价不仅体现在查询处理的策略上，还可能包括分布事务处理的管理策略。此外，在分布式环境中，查询处理分成全局查询和局部查询两个部分，其中对 I/O 的操作次数实际取决于操作方法。对联接操作、并操作不仅局部处理时有 I/O 费用，还涉及到全局查询时对不同场地间涉及二元操作的策略中的 I/O 代价因素。

专门讨论分布式查询中的传输代价也很有意义，因为传输代价很直观，一般来讲它是场地间数据传输量的函数，不象 I/O 代价还取决于完成操作的方法；同时，与分布式查询优化的特点相呼应，对于全局查询，讨论场地间存取策略的分布性时只考虑传输代价，而在决定每个场地上的局部存取策略时，再按集中式数据库方式来处理，这两个问题可以分开进行，而一般前者更为重要。

设一个查询执行的预期代价为 QC ，则在集中式中：

$$QC = \text{I/O 代价} + \text{CPU 代价}$$

在分布式中:

$$QC = \text{I/O 代价} + \text{CPU 代价} + \text{通信代价}$$

通信代价可用如下公式作粗略估算:

$$TC(X) = C_0 + C_1 \times X$$

其中, X 为数据的传输量, 通常以 bit 为单位计算; C_0 为两站点间通信初始化一次所花费的时间, 它由通信系统确定, 近似一个常数, 以秒为单位; C_1 为传输率 (传输速度的倒数), 即单位数据传输的时间, 单位是 b/s。